2014

Computational modeling of an economy using elements of artificial intelligence

Ekaterina Sinitskaya
Iowa State University

Computational modeling of an economy using elements of artificial intelligence

by

Ekaterina Sinitskaya

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:

Leigh S. Tesfatsion, Major Professor

Joydeep Bhattacharya

Steven M. Kautz

Peter F. Orazem

Sarah M. Ryan

Iowa State University

Ames, Iowa

2014

DEDICATION

I would like to dedicate this thesis to my husband Anton and to my son Michael without whose support and love I would not have been able to complete this work. I would also like to thank my friends and family for their loving guidance and financial assistance during the writing of this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to those who helped me with various aspects of conducting research and the writing of this thesis. First and foremost, Dr. Leigh S. Tesfatsion for her guidance, patience and support throughout this research and the writing of this thesis. Her insights and words of encouragement have often inspired me and renewed my hopes for completing my graduate education. I would also like to thank my committee members for their efforts and contributions to this work: Dr. Joydeep Bhattacharya, Dr. Steven M. Kautz, Dr. Peter F. Orazem, Dr. Sarah M. Ryan. I would additionally like to thank Dr. Harvey E. Lapan for his inspirational teaching style.

ABSTRACT

The goal of this dissertation was to develop tools for analyzing economic performance while agents were constrained to be constructively rational. To achieve this goal, firstly, tools for introducing forward-looking agents into agent-based frameworks were developed. These agents were shown to be a feasible alternative to the assumption of rational expectations, albeit with some limitations, as could be expected from any computational method. Several testing frameworks were also developed. Smaller ones were used to explore economic effects of decision procedures used by agents on macro- and micro-levels. A more advanced framework was formulated to facilitate the analysis of the interactions between institutional structures and macroeconomic policies. These frameworks were shown to be scalable and useful tools for the analysis of both micro-level decisions of agents and macroeconomic policies of central banks.

CHAPTER 1. INTRODUCTION

It is important to understand the origin and consequencies of economic crises, particularly now, when the world economy is remaining in a state of slow growth, despite all unconventional policies that authorities in different countries have been trying to implement since the financial crisis of 2008. The complexity of the most recent crises requires developing new tools that could help policy makers to understand possible side effects of various policies. Such tools could be designed in the agent-based macroeconomic paradigm.

The continuing development of computational facilities provides researches and government agencies with an access to vast computational resources. This development is finally allowing for much more complicated models to be employed. It is now possible to achieve a reasonable level of detail without incurring prohibiting computational costs. Besides that, modern computational resources make the problem of balancing the calculation time and the complexity of the model a much easier task. It is now getting possible to build models that not only have complicated institutional structures, but also include fairly advanced decisions procedures.

Thanks to these development, new questions arise that are worth investigating. Some of this question are: How should we model more sophisticated agents? What institutional features of the real economy should be included? This works begins to answer such questions.

The analysis starts with the introduction of a constructive rationality concept. It is suggested that this concept should be used as a modeling tool guiding efforts to incorporate more sophisticated decision procedures into the toolset of the agent-based macroeconomics. In the same chapter, an agent-based small scale macroeconomic framework is presented and forward-looking agents are introduced in this framework. These agents are shown to be a feasible alternative to the assumption of rational expectations, albeit with some limitations, as could be expected from any computational method.

In the next part of the dissertation, the problem of optimal belief structures is explored. The model was developed and tested that used different levels of Bayesian networks for the modeling of the belief structures.

Finally, a middle-scale macroeconomic model was developed that served as an analytical tool for the investigation of the effects of the central bank policies. A need to better understand interactions between institutional structures and macroeconomic policies is highlighted based on the analysis results produced by agent-based model.

In total, the models and the corresponding computer codes designed in this dissertation contribute to the understanding of the effects of constructive rationality of economic agents and the policy of the central bank on the economy.

CHAPTER 2. MACROECONOMIES AS CONSTRUCTIVELY RATIONAL GAMES

Real-world decision-makers are forced to be locally constructive, in the sense that their actions are constrained by the interaction networks, limited information, and computational capabilities at their disposal. This study poses the following question: Suppose utility-seeking consumers and profit-seeking firms in an otherwise standard dynamic macroeconomic model are required to be locally constructive decision-makers, unaided by the external imposition of global coordination conditions. What combinations of locally constructive decision rules result in good macroeconomic performance relative to a social planner benchmark model, and what are the game-theoretic properties of these decision-rule combinations? We begin our investigation of this question by specifying locally constructive decision rules for the consumers and firms that range from simple fixed behaviors to sophisticated approximate dynamic programming algorithms. We then use computational experiments to explore macroeconomic performance under alternative decision-rule combinations. A key finding is that simpler rules can outperform more sophisticated rules, but that forward-looking behavior coupled with a relatively long memory permitting past observations to inform current decision-making is critical for good performance.

2.1 Introduction

2.1.1 Study Overview

Decision-makers in real-world macroeconomies are necessarily limited to *locally constructive actions*, that is, to actions that can be implemented on the basis of their own interaction networks, limited information, and computational capabilities. In contrast, modern macroeconomic models typically impose coordination restrictions on the actions of decision-makers that are not locally constructive. Key examples include the global market clearing conditions and strong-form rational expectations postulates imposed in *dynamic stochastic general equilibrium (DSGE)* models.

These observations raise the following question. Suppose all actions within an otherwise standard DSGE model are required to be locally constructive, unaided by global coordination restrictions imposed by the modeler. What form could these locally constructive actions take to ensure good outcomes, not only for the individual participants but also for the macroeconomy as a whole?

This study addresses this question for a simplified version of the DSGE model developed by Smets and Wouters (2003) consisting of consumers and firms interacting over time in labor and goods markets. Each consumer desires to maximize his expected intertemporal (lifetime) utility subject to budget constraints, and each firm desires to maximize its expected intertemporal profit subject to technology constraints. However, in a departure from Smets and Wouters, the consumers and firms are restricted to *constructively rational decision procedures* in the following sense: the specification by these agents of their objective functions, decision domains, and decision rules mapping decision domains into decision selections must constitute locally constructive actions for these agents.

To investigate the implications of constructive rationality for the resulting *Dynamic Macroeconomic (DM) Game*, the decision domains for consumers and firms are first expressed in stationary form, as vectors of possible parameter selections. Each decision (parameter vector) maps into a sequence of parameterized supply and demand functions for current and future periods. Systematic computational experiments are then conducted to explore the implications of assuming that consumers and firms make successive selections from these decision domains in accordance with decision rules ranging from simple adaptation to sophisticated anticipatory learning. These decision rules include: (i) a reactive reinforcement learning method developed by Roth and Erev (1995) and Erev and Roth (1998) on the basis of findings from human-subject experiments; (ii) a forward-looking learning method developed by Watkins (1989), called Q-learning; (iii) a forward-looking rolling-horizon learning method (Alden and Smith (1992)); and (iv) an adaptive dynamic programming (ADP) learning method based on value-function approximation.

The key issue of interest is which decision-rule combinations come closest to achieving the benchmark optimal solution obtainable by a fully informed social planner. In particular, do the decision rules making relatively more sophisticated use of information tend to result in relatively higher welfare outcomes, either for the individual decision-rule users or for the economy at large? Since previous experimental findings have shown that minimally-informed traders using relatively unsophisticated decision rules can match or exceed the performance of better informed traders in some market contexts (Gode and Sunder (1993), Smith (2008)), the answer to this question is not obvious *a priori*. A related issue of interest is which (if any) decision-rule combinations constitute Nash equilibria and/or Pareto optimal solutions for the DM Game.

A key finding of this study is that good performance in the DM Game requires decision-makers to engage both in the exploitation of their current information and in searches for new information. Simpler rules can outperform more sophisticated decision rules, but only if the simpler rules entail forward-looking behavior coupled with a relatively long memory permitting past observations to inform current decision-making.

This study is organized as follows. Section 2.2 explains the basic structure of the DM Game together with its market and payment processes. Section 2.3 discusses the decision procedures implemented by the DM Game consumers and firms. Section 2.4 introduces and solves a social planner benchmark model as a benchmark of comparison for the simulation experiments. Section 2.5 describes the sensitivity design for simulation experiments, and Section 2.6 reports key simulation findings. Some technical implementation aspects are relegated to the Appendix, and the code is available at https://github.com/wilfeli/DMGameBasic.

2.1.2 Relationship to Previous Research

Numerous previous researchers, including Simon (1978), Dosi and Egidi (1991), Stiglitz (2002), Smith (2008), Howitt (2008), and Kahneman (2011), have emphasized the importance and complexity of modeling real-world decision-making procedures. Practitioners have also been interested in obtaining an improved understanding of these procedures; see, for example, a recent report (Trichet (2010)) by the President of the European Central Bank.

One possible approach permitting the systematic study of decision-making procedures is *Agent-based Computational Economics (ACE)*, the computational modeling of economic processes (including whole economies) as open-ended dynamic systems of interacting agents (Tesfatsion and Judd (2006), Tesfatsion (2014c)). Agents in ACE models can range from passive system entities with no cognitive function to active information-gathering decision-makers capable of sophisticated social and learning behaviors. The repeated interactions of these agents give rise to global regularities characterizing the system as a whole, which in turn affect agent interactions.

To date, however, ACE researchers have typically used decision procedures for macroeconomic agents that are not explicitly derived from underlying optimization problems. For example, Dawid et al. (2011), Oeffner (2008), Dosi et al. (2010), and Mandel et al.

(2010) directly model the behavior of consumers and firms using combinations of simple fixed and adaptive decision rules.[1]

In contrast, DSGE researchers typically assume that consumers and firms solve intertemporal utility and profit maximization problems; see, for example, Sbordone et al. (2010) and Tovar (2009). Yet, to avoid aggregation issues, DSGE researchers also typically assume the existence of representative consumer and firm agents with strong forms of rational expectations. This reliance on representative agents with rational expectations has been criticized on the grounds it prevents the study of learning and coordination issues critical for understanding the operation of real-world macroeconomies (Howitt (2012)).

A key point to stress here, however, is that agents in ACE models do not have to be restricted to reactive stimulus-response behavior; they can be modeled as forward-looking intertemporal optimizers.[2] Conversely, agents in DSGE models do not have to be modeled as optimizers with incredible information and computational capabilities; they can be modeled as learners reacting to experienced events.

Consequently, why not combine the best of these two approaches by examining constructively rational decision-making for economic agents with intertemporal goals? In particular, what forms (if any) of constructively rational decision-making by participants in macroeconomies result in good intertemporal outcomes, not only for the individual participants but also for the macroeconomy as a whole? The current study focuses on this issue.

A final note on terminology is in order. Our conception of a constructively rational decision procedure does not necessarily entail the pursuit of goals through the solution of optimization problems. Consequently, it differs from the concept of *procedural rationality*

[1]See Chen (2012) for a recent survey of ACE agent modeling, and see Tesfatsion (2014a) for extensive annotated pointers to ACE macroeconomic research.

[2]For an extensive collection of annotated pointers to research on learning algorithms for ACE agents, including approximate dynamic programming and other forward-looking methods for intertemporal optimization, see Tesfatsion (2014b).

introduced by Simon (1978)[p. 9], in which decision-making agents are assumed to pursue the most effective possible procedures for the choice of their actions, given their limited information and cognitive powers. Similarly, it differs from the concept of *constructivist rationality* introduced by Smith (2008)[p. 2], defined as "the deliberate use of reason to analyze and prescribe actions judged to be better than alternative feasible actions that might be chosen."

Rather, our conception permits *procedural uncertainty* (Dosi and Egidi (1991), Howitt (2008)), in the sense that decision-makers might be uncertain how to use their limited decision-making resources in an attempt to achieve their goals. In this case they might engage in a combined learning and decision-making process in an attempt to reduce their uncertainty about their world even as they attempt to survive and prosper within that world. Indeed, the operative question for a reader of this study is as follows: If you were to be suddenly transported into the DM Game as a consumer or firm, forced to implement your decisions in a locally constructive manner, what decision procedure would you use in an attempt to achieve your utility or profit goal?

2.2 The Dynamic Macroeconomic Game

2.2.1 Overview

This section develops a *Dynamic Macroeconomic (DM) Game*, a simplified version of the DSGE model developed by Smets and Wouters (2003) that will permit us to investigate the effects on micro and macro outcomes when consumers and firms use different decision procedures. A deliberate attempt has been made to ensure that the structure of the DM Game is similar to the structure of the Smets-Wouters DSGE model. However, the DM Game differs from this model in two critical ways: (i) absence of globally-imposed coordination conditions; and (ii) endogenous heterogeneity.

Regarding (i), in attempting to achieve their goals through participation in market processes, each consumer and firm in the DM Game is restricted to constructively rational decision procedures. As will be seen below, this requirement implies that events must proceed through historical time from cause to effect, with no non-causal looping permitted. In particular, the standard DSGE determination of market outcomes, in which labor and goods markets are simultaneously cleared at correct equilibrium prices with correct matching of buyers and sellers, with no risk to the traders, must be replaced by market processes permitting risky trades to proceed even if transactions are based on imperfectly informed demands and supplies.

Regarding (ii), heterogeneity among the DM Game consumers and among the DM Game firms arises endogenously over time from two sources. One source is that all of the decision procedures tested for consumers and firms in this study are adaptive procedures involving stochastic aspects in their implementations. A second source is the use of a stochastic rationing rule in the market clearing processes for labor and goods.[3]

The next subsection provides a big-picture understanding of the basic DM Game structure. The remaining subsections then explain in greater detail the market and payment processes in the DM Game, as well as the structure of the intertemporal optimization problems for consumers and firms. A detailed description of the particular locally-constructive decision procedures to be tested for the consumers and firms by means of computational experiments is given in the following Section 2.3.

[3]As detailed in Sections 2.3.2 and 2.3.3, reservation wages and prices are used to determine demand and supply functions in the DM Game. Agents thus suddenly enter or drop out of the labor and goods markets as the wage and price increase from 0, which induces discontinuities and flat portions in the aggregate demand and supply functions. In consequence, at the market clearing wage or price where the aggregate demand and supply curves cross each other, there can be too many units offered (or demanded) relative to demand (or supply). Random selection is used to determine which offers for units are used to clear demand in the case of excess supply and which demands for units are used to clear supply in the case of excess demand.

2.2.2 Basic DM Game Structure

As depicted in Fig. 2.1, the DM Game consists of a finite collection I of utility-seeking infinitely-lived consumers and a finite collection J of profit-seeking infinitely-lived corporate firms that interact in market and payment processes over discrete periods $t \geq 0$, where period $t = [t, t+1)$.

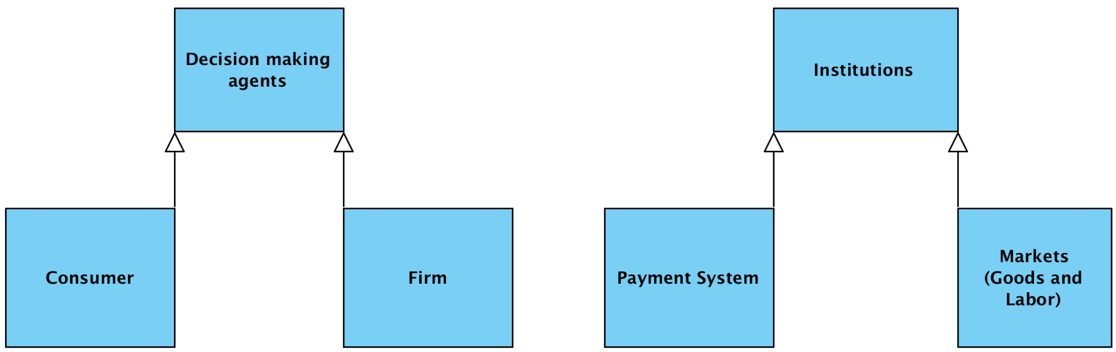

Figure 2.1 Decision-making agents and institutions for the DM Game

Each consumer and firm has an initial money balance at time 0, measured in book credit; and all subsequent payments and receipts take the form of changes in consumer and firm money balances. The consumers derive utility from leisure and from the consumption of a durable good q purchased from firms. The firms earn profits from the sale of good q to consumers, where q is produced by means of labor services purchased from consumers.

Both the labor market and the goods market are organized as competitive markets in which demands and supplies are matched to determine market-clearing prices and quantities. All firm profits are distributed back to consumers in the form of dividend payments. The goal of each consumer is to maximize his expected intertemporal utility subject to budget constraints, where this optimization problem is expressed in locally constructive terms. The goal of each firm is to maximize its expected intertemporal profits subject to technology constraints, where this optimization problem is expressed in locally constructive terms.

Each consumer at time 0 is structurally identical to each other consumer; that is, each consumer has the same initial money balance, human capital endowment, and intertemporal utility function. Also, each consumer owns an equal share of each firm, fixed through time, and hence receives the same steam of dividend payments. Similarly, each firm at time 0 is structurally identical to each other firm, meaning that each firm has the same initial money balance, goods stock, dividend allocation rule, and intertemporal profit function.

Market trades in the DM Game are risky in the following sense. In each period the labor market occurs prior to the goods market. Firms engage in forward contracting with consumers for labor services, and carry out goods production using these labor services, prior to the realization of actual goods demands. Firms thus risk bankruptcy if insufficient goods are sold to permit them to meet their wage obligations; and bankrupt firms must exit the DM Game economy. On the other hand, consumers risk non-payment for labor services rendered if firms become bankrupt. Since all goods demands must be backed by actual purchasing power, this can reduce the goods demands of the consumers in the next trading period, exacerbating firm cash-flow problems.[4]

A key question to be addressed is therefore as follows. Given the potential riskiness of market trading, and the restriction to locally constructive decision rules, is it worthwhile for the consumers and firms to use relatively sophisticated decision rules derived from intertemporal optimizations? Or should they instead proceed cautiously with simpler forms of decision rules based on incremental adaptations to past trading outcomes?

[4]For simplicity, this study assumes that consumer subsistence needs for goods are zero. Hence, the consumers do not face a risk of death by starvation if they are unable to purchase any goods.

2.2.3 Market and Payment Processes in the DM Game

All transactions in the DM Game are accompanied by corresponding payments, hence the payment system is an important underlying institution. For simplicity, this payment system is taken to be a simple clearing house that instantaneously clears each transaction. Although consumers and firms can carry forward savings in the form of money (book-credit), there is no banking system, hence no borrowing/lending opportunities and no interest paid on savings.

A consumer is not permitted to spend more than his current money holdings, hence all consumer demands for goods must be backed by actual purchasing power. A firm is declared bankrupt, and removed from the economy, if its current money holdings are insufficient to meet its wage-payment obligations to its workers.[5]

The consumers and firms use decision rules in each period t in an attempt to take actions that satisfy their intertemporal utility and profit goals. These actions consist of both labor and goods market decisions, such as whether or not to participate in these markets and what specific quantity and price terms to seek if they do. The consumers and firms receive feedback from the economy as a result of their period-t actions, and they update their decision rules on the basis of this feedback in preparation for period $t + 1$. This feedback includes market-clearing wages and prices for the period-t labor and goods markets, and their own private utility or profit outcomes as a result of their period-t market transactions.

[5]Any money held by a bankrupt firm is divided equally among its workers in partial fulfillment of its wage-payment obligations. However, goods stocks of bankrupt firms are assumed to be lost to the economy.

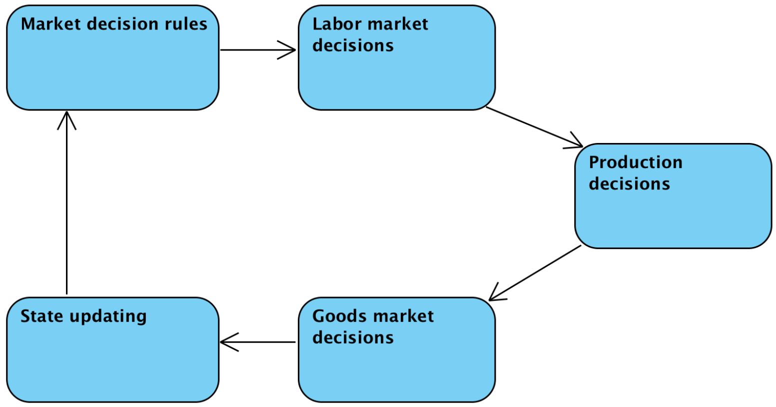

Figure 2.2 Sequential market decisions during a typical period t.

As depicted in Fig. 2.2, the labor market occurs before the goods market in each period t. Each consumer participating in the labor market submits a labor supply offer, and each firm participating in the labor market submits a labor demand bid. A labor market clearing solution is then calculated based on these offers and bids. This solution consists of a set of forward labor contracts (supply now, get paid later) that determine the amount of labor to be supplied now by each consumer to each firm, and the (common) wage to be paid later by the firms to the consumers for each unit of supplied labor.

After the close of the period-t labor market, the consumers perform labor for the firms in accordance with their forward labor contracts, which results in produced amounts of goods. Next, each consumer participating in the period-t goods market submits a goods demand bid, and each firm participating in the period-t goods market submits a goods supply offer. A goods market clearing solution is then calculated based on these bids and offers. This solution consists of a set of spot contracts that determine the amount of good to be received now by each consumer from each firm, and the (common) goods price to be paid now by the consumers to the firms for each unit of good received.

After the close of the period-t goods market, each firm proceeds to deliver goods to its customers, in return for goods payments, in accordance with its period-t goods market spot contracts. Each firm then settles its period-t wage-payment obligations to its workers as determined by its period-t forward labor contracts, if it has sufficient money holdings to cover these obligations. Otherwise, the firm is bankrupt and must exit the economy.

At the end of period t, each consumer calculates its period-t utility on the basis of its period-t consumption of leisure and goods. Also, each (non-bankrupt) firm calculates its period-t profit as its period-t goods-sales revenues minus its period-t wage payments. These period-t utility and profit outcomes are used by the consumers and firms to update their decision rules for period $t+1$.

A portion of any positive profits accrued by a firm during period t is distributed to the firm's consumer-owners as dividend payments at the end of period t. The wage and dividend payments received by a consumer from the firms at the end of period t, together with any other unspent monies held by the consumer at the end of period t, constitute the money balances of the consumer at the start of period $t+1$ to be used for goods purchases in period $t+1$.

This flow of events is illustrated in Fig. 2.3. Note the use of internal times t:1 through t:6 for events occurring within each period $t = [t, t+1)$.

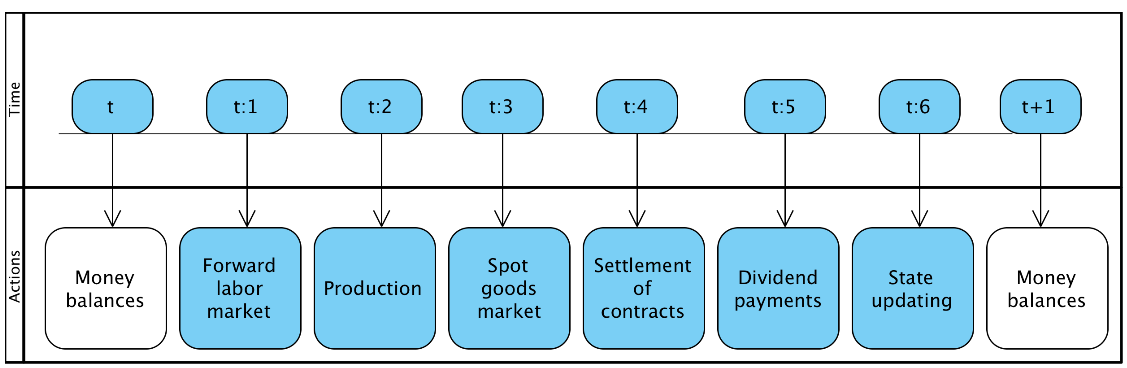

Figure 2.3 Flow of events during a typical period t.

As indicated in Fig. 2.3, the money balances held by consumers and firms at the end of period t (i.e., at time $t+1$) are determined by the money balances held by consumers and firms at the start of period t together with the additions and subtractions to these money balances arising from period-t market and dividend payments.

2.2.4 Consumer Constraints and Goals in the DM Game

Consumers in the DM Game are structurally identical. Each consumer i is endowed with the same initial positive money balance $M_{-1}^{c,0}$ (in book credit form) at the initial time 0. Each consumer i also has one unit of time in each period $t \geq 0$ that can be divided between labor services $l_{i,t:1}^{c}$ and leisure $[1 - l_{i,t:1}^{c}]$. For simplicity, it is assumed in this study that each consumer i in each period t devotes his one unit of time either all to labor or all to leisure.

Ignoring uncertainties (for the moment), the budget constraints faced by each consumer i in each period t take the following form:

$$s_{i,t:3} = M_{i,t-1}^{c} - p_{t:3}q_{i,t:3}^{c} \tag{2.1}$$

$$M_{i,t}^{c} = s_{i,t:3} + w_{t:4}l_{i,t:1}^{c} + div_{t:5}^{c} \tag{2.2}$$

$$s_{i,t:3},\ q_{i,t:3}^{c} \ \geq \ 0 \tag{2.3}$$

$$l_{i,t:1}^{c} \in \{0, 1\} \tag{2.4}$$

Here $M_{i,t-1}^{c}$ denotes consumer i's money balance at the start of period t, $p_{t:3}$ denotes the goods price determined in the goods market at time t:3 (same for all consumers), $q_{i,t:3}^{c}$ denotes the amount of good purchased by consumer i in the goods market at time t:3, $s_{i,t:3}$ denotes the savings of consumer i immediately subsequent to the goods market at time t:3, $w_{t:4}l_{i,t:1}^{c}$ denotes the actual wage payment received by consumer i at time t:4 arising from its forward labor contract cleared in the labor market at time t:1, and $div_{t:5}^{c}$ denotes the dividend payment (same for all consumers) received by consumer i at time t:5. The non-negativity constraint $s_{i,t:3} \geq 0$ ensures that consumer i's goods purchase $q_{i,t:3}^{c}$ is backed by actual purchasing power (money holdings).

The goal of each consumer i at the beginning of each period $t \geq 0$ is to maximize his expected intertemporal utility over periods $r \geq t$ subject to the budget constraints (2.1) through (2.4). If the labor service and consumption levels of consumer i in periods $r \geq t$ are given by $\left\{l^c_{i,r:1}, q^c_{i,r:3}\right\}_{r=t}^{\infty}$, then the intertemporal utility attained by consumer i over periods $r \geq t$ is given by

$$U_{i,t} = \sum_{r=t}^{\infty} \beta^{r-t} u\left(q^c_{i,r:3}, 1 - l^c_{i,r:1}\right) \ , \tag{2.5}$$

where $\beta \in (0,1)$ is a time-preference discount parameter.

In summary, as detailed above, the constraints and goals of the structurally-identical consumers in the DM Game depend on the specific settings for $(M^{c,0}_{-1}, u(\cdot), \beta)$. However, consumers do not know in advance the decision procedures in use by firms and other consumers, hence they do not know in advance the market-clearing values for future goods prices and wages nor the extent to which their own future goods demands and labor supplies will be fulfilled. How each consumer i might address this uncertainty through various alternative specifications for its own locally-constructive decision procedure will be explained in Section 2.3.

2.2.5 Firm Constraints and Goals in the DM Game

Firms in the DM Game are structurally identical. Each firm j is endowed with the same initial positive money balance $M^{f,0}_{-1}$ (in book credit form) and the same initial goods stock q^{stock}_{-1} at the initial time 0. Also, each firm j has the same stationary production function $q = F(l)$ for the production of good q using labor services l. Ignoring uncertainties (for the moment), the constraints faced by each firm j in each period t are derived as follows.

Let $q^{stock}_{j,t-1}$ denote firm j's inventory of goods at the beginning of period $t \geq 0$. Suppose firm j purchases labor services $l^f_{j,t:1}$ in the time-t:1 labor market and uses these labor services to produce a goods amount $q^f_{j,t:2} = F(l^f_{j,t:1})$ at time t:2. The goods amount $q^f_{j,t:3}$

that firm j sells in the time-t:3 goods market cannot exceed its goods inventory at the beginning of period t plus its goods production at time t:2:

$$q^{stock}_{j,t-1} \;+\; F(l^f_{j,t:1}) \;\geq\; q^f_{j,t:3} \tag{2.6}$$

Firm j's goods inventory $q^{stock}_{j,t}$ at the start of period $t+1$ is then determined from the following inventory accumulation equation:

$$q^{stock}_{j,t} \;=\; q^{stock}_{j,t-1} \;+\; F(l^f_{j,t:1}) \;-\; q^f_{j,t:3} \tag{2.7}$$

In addition, firm j must worry about avoiding bankruptcy, since bankrupt firms (i.e., firms unable to meet their wage obligations) must exit the DM Game economy. Consequently, firm j only distributes dividends in period t if its goods market revenues $p_{t:3}q^f_{j,t:3}$ earned at time t:3 exceed its wage obligations $w_{j,t:1}l^f_{j,t:1}$ incurred in the forward labor market at time t:1 for settlement at time t:4. Moreover, firm j limits its dividend distributions to its profits (if any). Specifically, firm j's total dividend payments $div^f_{j,t:5}$ at time t:5 are determined in accordance with the following allocation rule:

$$div^f_{j,t:5} = \begin{cases} \kappa^{div} \cdot \left[p_{t:3}q^f_{j,t:3} - w_{t:1}l^f_{j,t:1}\right] & \text{if } \; p_{t:3}q^f_{j,t:3} - w_{t:1}l^f_{j,t:1} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.8}$$

where $\kappa^{div} \in [0,1]$. Given (2.8), the no-bankruptcy condition for firm j in period t guaranteeing its period-t wage obligations can be fulfilled takes the form

$$M^f_{j,t-1} \;+\; p_{t:3}q^f_{j,t:3} \;-\; w_{t:1}l^f_{j,t:1} \;\geq\; 0 \tag{2.9}$$

The money balance $M^f_{j,t}$ held by a non-bankrupt firm j at the end of period t (i.e., at the start of period $t+1$) is determined by the money balance $M^f_{j,t-1}$ held by firm j at the start of period t adjusted to reflect firm j's market activities and dividend payments during period t, as follows:

$$M^f_{j,t} \;=\; M^f_{j,t-1} + p_{t:3}q^f_{j,t:3} - w_{t:1}l^f_{j,t:1} - div^f_{j,t:5} \tag{2.10}$$

Finally, the following non-negativity restrictions on firm j's labor service demand $l^f_{j,t:1}$ at time t:1 and goods supply $q^f_{j,t:3}$ at time t:3 must be satisfied for physical meaningfulness:

$$l^f_{j,t:1},\ q^f_{j,t:3} \quad \geq \quad 0 \tag{2.11}$$

The goal of each firm j at the beginning of each period $t \geq 0$ is to maximize its expected intertemporal utility over periods $r \geq t$ subject to the technological and feasibility constraints (2.6) through (2.11). For any given sequence $\left\{w_{r:1}, l^f_{j,r:1}, p_{r:3}, q^f_{j,r:3}\right\}_{r=t}^{\infty}$ of wage levels, labor service purchases, goods prices, and goods purchases for periods $r \geq t$, the intertemporal profit attained by firm j over periods $r \geq t$ is given by

$$\Pi_{j,t} = \sum_{r=t}^{\infty} \mu^{r-t} \left[p_{r:3} q^f_{j,r:3} - w_{r:1} l^f_{j,r:1}\right] \tag{2.12}$$

where $\mu \in (0,1)$ is a time-preference discount parameter.

In summary, as detailed above, the constraints and goals of the structurally-identical firms in the DM Game depend on the specific settings for $(M^{f,0}_{-1}, q^{stock}_{-1}, F(\cdot), \mu, \kappa^{div})$. However, firms do not know in advance the decision procedures in use by consumers and other firms, hence they do not know in advance the market-clearing values for wages and goods prices nor the extent to which their own future labor supplies and goods demands will be fulfilled. How each firm j might address this uncertainty through various alternative specifications for its own locally-constructive decision procedure will be explained in the following Section 2.3.

2.3 Locally-Constructive Decision Procedures

2.3.1 Overview of Decision Procedures

The locally-constructive decision procedures to be tested for consumers and firms in the DM Game are processes for the adaptive determination of demand bids and supply offers for the labor and goods markets in each successive period t. The specification of these decision procedures is divided into three steps, as follows.

First, decision domains are specified for consumers and firms that consist of possible selections of "tuning" parameters for demand and supply functions. To permit more meaningful comparisons among decision procedures, the decision domain for each consumer at the beginning of each period t is specified as a cross-product D^c of finite sets, the same for each consumer. Similarly, the decision domain for each firm at the beginning of each period t is specified as a cross-product D^f of finite sets, the same for each firm.

Second, state-conditioned transformation functions are specified for consumers and firms. The state of a consumer or firm at any time t consists of the time-t physical attributes, information, and beliefs of this agent. The transformation function for each consumer at the beginning of each period $t \geq 0$ maps each of his possible decisions d^c in D^c into a collection of labor supply and goods demand functions for periods $r \geq t$, parameterized by d^c, and conditional on the consumer's time-t state. Similarly, the transformation function for each firm at the beginning of each period $t \geq 0$ maps each of its possible decisions d^f in D^f into a collection of labor demand and goods supply functions for periods $r \geq t$, parameterized by d^f, and conditional on this firm's time-t state.

Third, *Reactive Learner (RL)*, *Forward-looking Learner (FL)*, and *Explicit Optimizer (EO)* decision rules are specified for each consumer and firm that determine how this agent selects decisions from its decision domain in each period t. These three types of decision procedures cover a range of decision-making behaviors roughly ordered from less to more sophisticated with regard to information utilization, expectation formation, and forward-looking behavior. A summary description of these decision-maker types is given in Table 2.1.

Table 2.1 Types of decision procedures for consumers and firms in the DM Game.

Agent	Decision-Maker Type	Decision Procedure Description
Consumer	Reactive Learner (RL)	Adaptively updates decisions in response to realized utility outcomes
	Forward-Looking Learner (FL)	Uses Q-learning in an attempt to maximize expected intertemporal utility
	Explicit Optimizer (EO)	Maximizes expected intertemporal utility using adaptively updated probabilities
Firm	Reactive Learner (RL)	Adaptively updates decisions in response to realized profit outcomes
	Forward-Looking Learner (FL)	Uses Q-learning in an attempt to maximize expected intertemporal profit
	Explicit Optimizer (EO)	Maximizes expected intertemporal profit using adaptively updated probabilities

The construction of the decision domains and the state-conditioned transformation functions for consumers and firms is explained more carefully in Sections 2.3.2 and 2.3.3. Sections 2.3.4 through 2.3.6 then describe the decision rules used to select decisions from these decision domains for each of the three types of decision-makers RL, FL, and EO listed in Table 2.1.

2.3.2 Decision Domain and Transformation Function for Consumers

The decision domain D^c for each consumer i is given by a cross-product of finite sets having the form

$$D^c \;=\; L^c \otimes \Omega \otimes \Theta \tag{2.13}$$

where:

- $L^c = \{0, 1\}$
- the elements of $\Omega = \{\omega_1, \ldots, \omega_G\}$ satisfy $0 < \omega_1 < \ldots < \omega_G$
- the elements of $\Theta = \{\theta_1, \ldots, \theta_H\}$ satisfy $0 \leq \theta_1 < \ldots < \theta_H \leq 1$

Consumer i selects a decision $d^c = (l^c, \omega, \theta)$ from D^c at each time $t \geq 0$ by means of its particular RL, FL, or EO decision rule. The selection of d at time t is then transformed into a sequence $\mathbf{TR}^c_{i,t}(d)$ of labor supply and goods demand functions $(l^c_{i,r:1}(w, d, t), q^c_{i,r:3}(p, d, t))_{r \geq t}$, parameterized by d and conditional on consumer i's time-t state.

Specifically, the labor supply $l^c_{i,r:1}(w, d, t)$ as a function of the time-r:1 labor market wage w is determined as follows. If $l^c = 0$, then $l^c_{i,r:1}(w, d, t) = 0$ for all w, meaning that consumer i does not plan to participate in the time-r:1 labor market. On the other hand, if $l^c = 1$, the *reservation wage* of consumer i for the time-r:1 labor market is given by

$$w^c_{i,r:1}(d, t) \;=\; \omega \cdot E_{i,t}[w_{r:1}] \tag{2.14}$$

where $E_{i,t}[w_{r:1}]$ denotes the time-r:1 labor market wage expected by consumer i, based on his state at time t. If $w < w^c_{i,r:1}(d, t)$, then $l^c_{i,r:1}(w, d, t) = 0$, meaning that consumer i does not plan to participate in the time-r:1 labor market at this labor market wage. On the other hand, if $w \geq w^c_{i,r:1}(d, t)$, then $l^c_{i,r:1}(w, d, t) = 1$, meaning that consumer i plans to offer his 1 unit of labor service into the time-r:1 labor market at this labor market wage.

Also, the goods demand $q^c_{i,r:3}(p,d,t)$ as a function of the time-r:3 goods market price p takes the form

$$p \cdot q^c_{i,r:3}(p,d,t) \;=\; \theta \cdot M^c_{i,r-1} \tag{2.15}$$

Thus, consumer i plans in period t to spend a fraction θ of his time-r money balance $M^c_{i,r-1}$ on consumption goods at time r:3, and he specifies his time-r:3 goods demand as a function of the time-r:3 market price p in accordance with this plan. Note that $M^c_{i,r-1}$ will be known to consumer i at time r, prior to the opening of the goods market at time r:3.[6]

The decision domain D^c depends on the grid specifications for Ω and Θ; these grid specifications are explained in Appendix A.1. The transformation function $\mathbf{TR}^c_{i,t}$ depends on the wage expectation in (2.14). The method used by consumers to form and update their wage expectations is explained in Appendix A.2.

2.3.3 Decision Domain and Transformation Function for Firms

The decision domain D^f for each firm j is given by a cross-product of finite sets having the form

$$D^f \;=\; L^f \otimes \Gamma \otimes \Lambda \otimes \Psi \tag{2.16}$$

where:

- the elements of $L^f = \{l^f_1, \ldots, l^f_L\}$ satisfy $0 \leq l^f_1 < \ldots < l^f_L\}$
- the elements of $\Gamma = \{\gamma_1, \ldots, \gamma_M\}$ satisfy $0 < \gamma_1 < \ldots < \gamma_M$
- the elements of $\Lambda = \{\lambda_1, \ldots, \lambda_N\}$ satisfy $0 < \lambda_1 < \ldots < \lambda_N$
- the elements of $\Psi = \{\psi_1, \ldots, \psi_R\}$ satisfy $0 \leq \psi_1 < \ldots < \psi_R \leq 1$

[6]Recall that consumer i receives no money payments between time r (the beginning of period r) and the settlement of labor market contracts at time r:4. Thus, consumer i's purchases in the time-r:3 goods market cannot exceed his money balance $M^c_{i,r-1}$ at time r.

Firm j selects a decision $d = (l^f, \gamma, \lambda, \psi)$ from D^f at each time $t \geq 0$ by means of its particular RL, FL, or EO decision rule. The selection of d at time t is then transformed into a sequence $\mathbf{TR}^f_{j,t}(d)$ of labor demand and goods supply functions $(l^f_{j,r:1}(w,d,t), q^f_{j,r:3}(p,d,t))_{r \geq t}$, parameterized by d and conditional on firm j's time-t state.

Specifically, the labor demand $l^f_{j,r:1}(w,d,t)$ as a function of the time-r:1 labor market wage w is determined as follows. If $l^f = 0$, then $l^f_{j,r:q}(w,d,t) = 0$ for all w, meaning that firm j does not plan to participate in the time-r:1 labor market. If $l^f > 0$, the *reservation wage* of firm j for the time-r:1 labor market is given by

$$w^f_{j,r:1}(d,t) \;=\; \gamma \cdot E_{j,t}[w_{r:1}] \tag{2.17}$$

where $E_{j,t}[w_{r:1}]$ denotes the time-r:1 labor market wage expected by firm j, based on its state at time t. If $w > w^f_{j,r:1}(d,t)$, then $l^f_{j,r:1}(w,d,t) = 0$, meaning that firm j does not plan to participate in the time-r:1 labor market at this labor market wage. On the other hand, if $w \leq w^f_{j,r:1}(d,t)$, then $l^f_{j,r:1}(w,d,t) = l^f$, meaning that firm j plans to demand l^f units of labor in the time-r:1 labor market at this labor market wage.

Also, the goods supply $q^f_{j,r:3}(p,d,t)$ as a function of the time-r:3 goods market price p is determined as follows. The *reservation goods price* of firm j for the time r:3 goods market is given by

$$p^f_{j,r:3}(d,t) = \lambda \cdot E_{j,t}[p_{r:3}] \tag{2.18}$$

where $E_{j,t}[p_{r:3}]$ denotes the time-r:3 goods market price expected by firm j, based on its state at time t. If $p < p^f_{j,r:3}(d,t)$, then $q^f_{j,r:3}(p,d,t) = 0$, meaning that firm j does not plan to participate in the time-r:3 goods market at this goods market price. On the other hand, if $p \geq p^f_{j,r:3}(d,t)$, then

$$q^f_{j,r:3}(p,d,t) \;=\; \psi \cdot q^{stock}_{j,r:2} \tag{2.19}$$

That is, firm j plans to supply a fraction ψ of its time-r:2 goods stock into the time-r:3 goods market at the goods market price p. Note that $q^{stock}_{j,r:2}$ will be known to firm j at time r:2, prior to the opening of the goods market at time r:3.

The decision domain D^f depends on the grid specifications for L^f, Γ, Λ, and Ψ; these grid specifications are explained in Appendix A.1. The transformation function $\mathbf{TR}^f_{j,t}$ depends on the wage expectation in (2.17) and the price expectation in (2.18). The method used by firms to form and update their wage and price expectations is explained in Appendix A.2.

2.3.4 RL Decision Rule for Consumers and Firms

Reinforcement learning embodies the basic common-sense principle that the propensity to select relatively good decisions should be reinforced and the propensity to select relatively poor decisions should be discouraged. Immediate rewards flowing from decisions are typically used to update the propensities for choosing these decisions in an appropriate up or down direction.

The RL decision rule for consumers and firms in the DM Game is a reinforcement learning method originally developed by Roth and Erev (1995) and Erev and Roth (1998) and subsequently modified by Nicolaisen et al. (2001). This method is “reactive” in the sense that it asks the following backward-looking question: Given past events, what decision should I make now?

For the DM Game, the immediate reward $R^c_i(d,t)$ received by a consumer i as a result of selecting a decision d in D^c at the beginning of any period t is taken to be consumer i's realized period-t utility. Similarly, the immediate reward $R^f_j(d,t)$ received by a firm j as a result of selecting a decision d in D^f at the beginning of any period t is taken to be firm j's realized period-t profit.

Below we explain the RL decision rule for an arbitrary decision-maker v who selects a decision d from a finite decision domain D in each period t, receiving an immediate reward $R(d,t)$, where v could represent either a consumer or a firm in the DM Game economy. Let the finite cardinality of D be denoted by $\mathcal{D}$, and let the elements of D be indexed by $d = 1,\ldots,\mathcal{D}$.

Suppose it is the beginning of the initial period 0, prior to decision selection, and suppose decision-maker v must select a decision from its decision domain D for period 0. Suppose the *initial propensity* of v to select decision d in D at time 0 is exogenously given by $q(d,0)$ for $d = 1,\ldots,\mathcal{D}$. Let the vector of these initial propensities be denoted by $\mathbf{q}(0) = (q(1,0),\ldots,q(\mathcal{D},0))$.

Now suppose it is the beginning of any period $t \geq 0$, prior to decision selection, and suppose the current propensity of decision-maker v to select decision d in D is given by $q(d,t)$ for $d = 1,\ldots,\mathcal{D}$. The *choice probabilities* that v uses to select a decision for period t are then constructed from these propensities as follows:

$$\text{Prob}(d,t) = \frac{\exp(q(d,t)/C)}{\sum_{k=1}^{\mathcal{D}} \exp(q(k,t)/C)} \ , \ d = 1,\ldots,\mathcal{D} \tag{2.20}$$

In (2.20), C is a *cooling parameter* that affects the degree to which v makes use of propensity values in determining his choice probabilities. As $C \to \infty$, then $\text{Prob}_d(t) \to 1/\mathcal{D}$, so that in the limit v pays no attention to propensity values in forming his choice probabilities. On the other hand, as $C \to 0$, the choice probabilities (2.20) become increasingly peaked over the particular decisions d having the highest propensity values $q(d,t)$, thereby increasing the probability that these decisions will be chosen by v.

At the end of period t, the current propensity $q(d,t)$ that decision-maker v associates with each decision d in D is updated in accordance with the following rule. Let d_t in D denote the decision that v *actually* selected and implemented during period t. Also, let $R(d_t,t)$ denote the reward attained by v at the end of period t as a result of the implementation of d_t. Then, for each decision d in D,

$$q(d,t+1) \quad = \quad [1-r]q(d,t) \ + \ \text{Response}(d,t) \ \ , \tag{2.21}$$

where

$$\text{Response}(d,t) = \begin{cases} [1-e]\cdot R(d_t,t) & \text{if } \ d = d_t \\ \\ e \cdot q(d,t)/[\mathcal{D}-1] & \text{if } \ d \neq d_t, \end{cases} \tag{2.22}$$

and $d \neq d_t$ implies $\mathcal{D} \geq 2$. The *recency parameter* $r \in [0, 1]$ appearing in (2.21) controls the relative weighting of past versus current rewards in the updating of the propensities. The *experimentation parameter* $e \in [0, 1)$ appearing in (2.22) permits reinforcement to spill over from a chosen decision to other decisions to encourage experimentation with various decisions in the early stages of the learning process.

In summary, the RL decision rule is fully characterized once values are specified for $(\mathcal{D}, \mathbf{q}(0), C, e, r)$. Note that the RL decision rule is well-defined for any decision domain with finite cardinality $\mathcal{D}$; the exact form of the decisions constituting this decision domain is irrelevant. Note, also, that the decision-maker does not need to know his reward function; the RL decision rule only makes use of realized rewards, not potential rewards. The versatility and low-information requirements of the RL decision rule, together with its demonstrated robust performance in diverse situations, have led to its widespread use in learning applications.

2.3.5 FL Decision Rule for Consumers and Firms

The FL decision rule for consumers and firms in the DM Game is a "greedy" variant of the Q-learning algorithm developed by Watkins (1989) that permits decisions to be taken in accordance with dynamic programming policy functions in approximate form. The FL decision rule is "forward looking" in the sense that it asks the following anticipatory question: If I make this decision now, what will happen in the future?

The key conceptual construct underlying Q-learning (and dynamic programming in general) for a decision-maker v is the *value function* $V_t(x)$, defined to be the optimum total reward that can be obtained by v, starting at time t in state x. Below we provide an intuitive derivation of ϵ-greedy Q-learning as a policy-function approximation method, without consideration of technical details regarding the existence and uniqueness of optimal solutions.

Suppose a decision-maker v is currently in state x at some current time t. Suppose v implements a decision d, obtains an immediate reward $R_t(x,d)$, and transits to a new state $x' = S_t(x,d)$. Then the best that v can do, starting from time $t+1$, is $V_{t+1}(x')$. Consequently, the best v can do, starting from time t, is

$$V_t(x) = \max_d \left[R_t(x,d) + V_{t+1}(S_t(x,d))\right] \tag{2.23}$$

Finally, let π^* denote the *optimal policy function* giving the optimal decision d^* in (2.23) as a function $d^* = \pi^*(t,x)$ of the current time t and state x. Then (2.23) can equivalently be written as

$$V_t(x) = \left[R_t(x,\pi^*(t,x)) + V_{t+1}(S_t(x,\pi^*(t,x)))\right] \tag{2.24}$$

The recursive relationships (2.23) and (2.24) provide simple deterministic illustrations of Richard Bellman's celebrated *principle of optimality*.[7] As detailed in Powell (2011, 2014), one practical difficulty is how to compute the value function $V_t(x)$ and/or the optimal policy function π^*. Another practical difficulty is that the reward function $R_t(x,d)$ and/or the state transition function $S_t(x,d)$ might not be known; for example, they could depend on the unknown decisions of other agents in the system.

The Q-learning method provides a way to implement decisions in approximate accordance with the optimal policy function π^*, assuming the decision horizon is infinite and the reward, state transition, and value functions are independent of time. Below we provide a general description of this method.

For each state x and decision d, define

$$Q(x,d) = \left[R(x,d) + V(S(x,d))\right] \tag{2.25}$$

[7]Stochastic versions of the principle of optimality can be obtained by assuming R_t and/or S_t are influenced in each period t by the realization ω_t of a random event from a well-defined probability space $(\Omega, \mathcal{F}, \mathcal{P})$. An expectation (with respect to ω_t) is then taken of the bracketed term on the right-hand side of (2.23) prior to undertaking the maximization. More complex stochastic variants are obtained if the probability space for ω_t depends on the time t, the time-t state, and/or the decision-maker's time-t decision. See Powell (2014) for details.

If the Q-values in (2.25) can be learned, then the optimal policy function π^* is determined as follows: For any state x,

$$\pi^*(x) = \arg\max_d Q(x, d) \tag{2.26}$$

Hence, the learning of the Q-values in (2.25) avoids the need for separate learning or knowledge of the reward, state transition, and value functions.

In its simplest form, Q-learning uses the following iterative procedure to determine estimates $\widehat{Q(x,d)}$ for the Q-values $Q(x, d)$ in (2.25) conditional on a user-specified learning rate α and a user-specified discount factor γ:

Step 1: Initialize $\widehat{Q(x,d)}$ to a random value for each possible state x and decision d.

Step 2: Observe an actual state x'.

Step 3: Pick a decision d' and implement it.

Step 4: Observe the next state x'' and the next reward R''.

Step 5: Update the estimate $\widehat{Q(x',d')}$ as follows:

$$\widehat{Q(x',d')} \quad \leftarrow \quad [1-\alpha]\widehat{Q(x',d')} + \alpha\left[R'' + \gamma \max_d \widehat{Q(x'',d)}\right] \tag{2.27}$$

Step 6: Loop back to Step 2 and repeat.

The above procedure does not specify how the decision in Step 3 is to be picked. Let ϵ be any number in $(0, 1)$. The ϵ-greedy variant of Q-learning replaces the above Step 3 with an alternative Step $3'$ incorporating a specific decision selection process that accommodates two goals: (i) Exploit current information for maximum possible current gain; and (ii) seek new information to improve opportunities for future gains. This decision selection process is as follows: With probability ϵ the decision-maker v in Step $3'$ experiments by selecting a random decision d'. However, with probability $[1 - \epsilon]$ the decision-maker v instead "greedily" chooses a decision d^* that maximizes the current estimator $\widehat{Q(x', d)}$ for $Q(x', d)$.

In summary, the ϵ-greedy Q-learning method for a decision-maker v is fully characterized once values are specified for the initial Q-value estimates $\widehat{Q(x, d)}$ and the three parameters $(\gamma, \epsilon, \alpha)$.

2.3.6 EO Decision Rules for Consumers and Firms

Each EO agent (consumer or firm) at the beginning of each period $t \geq 0$ attempts to maximize an explicit expression for their expected reward (intertemporal utility or profit) over current and future periods $r \geq t$, subject to constraints. The EO agents use an "open-loop/closed-loop" optimization approach in the following sense: They undertake their maximization problems in each period t conditional on updated state information, yet in these maximizations they ignore the fact that they will re-optimize their period-t decision selections at the beginning of each future period $r > t$. They also ignore that rationing can occur on the margin in the market clearing processes.

Specifically, at the beginning of each period $t \geq 0$ an EO consumer i selects a decision d in D^c that maximizes his expected intertemporal utility over current and future periods $r \geq t$. In this maximization, consumer i makes use of the transformation function $\mathbf{TR}^c_{i,t}(d)$ detailed in Section 2.3.2 to map each possible decision d in D^c at time t into a collection of current and future labor supply and goods demand functions $(l^c_{i,r:1}(w, d, t), q^c_{i,r:3}(p, d, t))_{r \geq t}$.

Formally stated, an EO consumer i's maximization problem at the beginning of each period $t \geq 0$ takes the following form:

$$\max_{d \in D^c} E_{i,t} U_t(\mathbf{TR}^c_{i,t}(d), \mathbf{w}_{t:1}, \mathbf{p}_{t:3}) \tag{2.28}$$

subject to the budget and feasibility constraints (2.1) through (2.4) dependent on

$$\mathbf{w}_{t:1} = (w_{r:1})_{r=t}^{\infty} \tag{2.29}$$

$$\mathbf{p}_{t:3} = (p_{r:3})_{r=t}^{\infty} \tag{2.30}$$

$$\mathbf{div}_{t:5} = (div_{r:5})_{r=t}^{\infty} \tag{2.31}$$

where

$$U_t(\mathbf{TR}_{i,t}^c(d), \mathbf{w}_{t:1}, \mathbf{p}_{t:3}) = \sum_{r=t}^{\infty} \beta^{r-t} \left[u\left(q_{i,r:3}^c(p_{r:3}, d, t), 1 - l_{i,r:1}^c(w_{r:1}, d, t) \right) \right] \tag{2.32}$$

Similarly, an EO firm j's maximization problem at the beginning of each period $t \geq 0$ takes the following form:

$$\max_{d \in D^f} E_{j,t} \Pi_t(\mathbf{TR}_{j,t}^f(d), \mathbf{w}_{t:1}, \mathbf{p}_{t:3}) \tag{2.33}$$

subject to the technological and feasibility constraints (2.6) through (2.11) dependent on $\mathbf{w}_{t:1}$ and $\mathbf{p}_{t:3}$, defined as in (2.29) and (2.30), where

$$\Pi_t(\mathbf{TR}_{j,t}^f(d), \mathbf{w}_{t:1}, \mathbf{p}_{t:3}) = \sum_{r=t}^{\infty} \mu^{r-t} \left[p_{r:3} q_{j,r:3}^f(p_{r:3}, d, t) - w_{r:1} l_{j,r:1}^f(w_{r:1}, d, t) \right] \tag{2.34}$$

As explained in Appendix A.2, the expectations in the maximization problems (2.28) and (2.33) for each period t are based on estimated probability distributions for future labor market wages, future goods market prices, and future dividend payments (for consumers), conditional on the states of consumer i and firm j at time t.

As explained in Appendix A.3, approximate solutions for the maximization problems (2.28) and (2.33) are derived using two different approaches. Briefly summarized, the first approach, referred to as *EO Adaptive Dynamic Programming (EO-ADP)*, derives an approximate solution in each period t by solving a stochastic dynamic programming recurrence relation, assuming a basis-function approximation for the value function. The second approach, referred to as *EO Finite Horizon (EO-FH)*, replaces the infinite planning horizon in each period t with a finite planning horizon of length T, called the *forecasting horizon*, and then derives an approximate solution by means of direct search across the decision domain.

2.4 Social Planner Benchmark Model

The main source of uncertainty for each consumer and firm in the DM Game is *behavioral uncertainty*, meaning uncertainty concerning the decision procedures used by other consumers and firms. The only other source of uncertainty is the use of a random rationing rule in the labor and goods markets to determine which demanders receive goods or services in excess demand conditions and which suppliers sell goods or services in excess supply conditions; cf. footnote 3. There are no external shocks to the DM Game economy.

Both sources of uncertainty for the DM Game disappear if market decision-making by consumers and firms is replaced by a social planner who maximizes the intertemporal utility of a representative consumer i subject only to technological feasibility constraints, conditional on the restriction that the structurally-identical consumers must all be treated alike and the structurally-identical firms must all be treated alike. The resulting model, hereafter referred to as the *Social Planner (SP) Benchmark Model*, is introduced here in order to have a benchmark of comparison for the DM-Game simulation findings reported in Section 2.6.

Specifically, suppose the number I of DM-Game consumers and the number J of DM-Game firms are arbitrary positive integers, and let $q_{-1}^{stock} \geq 0$ denote the exogenously given goods stock of each firm at the beginning of period 0. We consider a social planner who solves the following social welfare optimization problem at time 0 on behalf of the representative DM-Game consumer:[8]

$$\max \sum_{t=0}^{\infty} \beta^t u(q_{t:3}^c, 1 - l_{t:1}^c) \tag{2.35}$$

with respect to $\{l_{t:1}^c, q_{t:3}^c\}_{t=0}^{\infty}$, subject to the following constraints for each $t \geq 0$:

$$J \cdot q_t^{stock} \;=\; J \cdot q_{t-1}^{stock} \;+\; J \cdot F(l_{t:1}^f) \;-\; I \cdot q_{t:3}^c \tag{2.36}$$

[8]Given the exponential form of the discount factor in (2.35), the social planner would exhibit time consistency, meaning that re-optimization in successive periods would not result in any deviation from the optimal solution determined at time 0.

$$l^f_{t:1} \;=\; \frac{I \cdot l^c_{t:1}}{J}$$

$$0 \;\leq\; q^{stock}_t,\; q^c_{t:3}$$

$$l^c_{t:1} \in \{0,1\}$$

To obtain a concrete SP Benchmark Model solution, we assume that the utility function $u(\cdot)$ in (2.35) takes the form

$$u(q, 1-l) \;=\; \delta^c_0 \cdot \ln\left(b(q) \;+\; q\right) + \delta^c_1 \cdot [1-l] \tag{2.37}$$

where[9]

$$b(q) = \begin{cases} 1.0 & \text{if} \;\; q > 0 \\ b \in (0,1) & \text{if} \;\; q = 0 \end{cases} \tag{2.38}$$

Also, the production function $F(\cdot)$ in (2.36) is assumed to take the form

$$F\left(l\right) \;=\; \delta^f_0 l^{\delta^f_1} \tag{2.39}$$

We further assume that the values specified for the parameters appearing in this SP Benchmark Model are as listed in Table 2.2. Finally, for each $t \geq -1$ we let

$$s^{stock}_t \;\equiv\; \frac{J \cdot q^{stock}_t}{I} \tag{2.40}$$

denote the per-consumer amount of goods stock carried forward from period t to period $t+1$.

[9]In order to permit consumers to constructively compare consequences for failure to participate in the goods market, the valuation they place on failure to participate needs to be finite. As will be seen in Section 2.6, the advantage of introducing the discontinuous valuation function $b(q)$ in (2.38) is that a consumer's utility takes on a negative value only if he fails to participate in the goods market, thus providing an easily detected signal of this non-participation.

Table 2.2 Maintained parameter values for the SP Benchmark Model and the DM Game

Parameter	Value
q_{-1}^{stock}	0.0
β	0.95
δ_0^c	3.0
δ_1^c	0.5
b	0.5
δ_0^f	1.0
δ_1^f	1.0

Given these concrete specifications, the SP Benchmark Model (2.35) can be expressed in the following reduced representative-consumer form:

$$\max \sum_{t=0}^{\infty} \beta^t \Big[3.0 \cdot \ln(b(q_{t:e}^c) + q_{t:3}^c) + 0.5 \cdot (1 - l_{t:1}^c)\Big] \tag{2.41}$$

with respect to $\{l_{t:1}^c, q_{t:3}^c\}_{t=0}^{\infty}$, subject to the following constraints for each $t \geq 0$:

$$s_t^{stock} \;=\; s_{t-1}^{stock} \;+\; l_{t:1}^c \;-\; q_{t:3}^c$$

$$0 \;\leq\; s_t^{stock},\; q_{t:3}^c$$

$$l_{t:1}^c \in \{0, 1\}$$

$$s_{-1}^{stock} = 0 \tag{2.42}$$

The solution of the reduced SP Benchmark Model (2.41) is a full-employment solution with $l_{t:1}^c = q_{t:3}^c = 1$ and $s_t^{stock} = 0$ for all $t \geq 0$. The proof, by induction, is provided in Appendix A.4.

Given this optimal solution, the representative consumer attains the stationary per-period utility level

$$u(1,0) = [3.0 \cdot \ln(2)] \approx 2.08 \tag{2.43}$$

and the intertemporal utility level

$$\sum_{t=0}^{\infty} \beta^t u(1,0) \;=\; \sum_{t=0}^{\infty} \beta^t 3.0 \cdot \ln(2) \;=\; 3.0 \cdot \ln(2) \frac{1}{1-\beta} \;\approx\; 41.59 \tag{2.44}$$

Note that the smallest single-period utility outcome that a representative consumer can feasibly attain under the SP Benchmark Model assumptions is $u(0,0) = 3.0 \cdot \ln(0.5) \approx -2.08$.

2.5 Sensitivity Design

2.5.1 Design Overview

The main focus of this study is the degree to which consumers in the DM Game economy are able to attain the one-period and intertemporal utility levels (2.43) and (2.44) achieved by the representative consumer in the SP Benchmark Model when the DM Game consumers and firms use different combinations of constructively-rational decision rules. The tested combinations of decision rules are displayed in Table 2.3.

Table 2.3 Tested combinations of constructively-rational decision rules (case numbers)

	C:RL	C:FL	C:EO-FH	C:EO-ADP
F:RL	1–10	21	31	39
F:FL	22	11–20	32	40
F:EO-FH	33	34	23–30	41
F:EO-ADP	42	43	44	35–38

For each of the 44 cases in Table 2.3, simulations were conducted for a range of values for a subset of parameters, hereafter referred to as the *treatment factors* for the case, while maintaining fixed values for all other parameters. For each tested combination of values for the treatment factors, the number of runs was set at NRuns = 10, using ten seed values for the random number generator.[10] The length of each run was set to LRun = 1000 periods. To reduce dependence on transient effects, outcomes from the first LOmit = 50 periods in each run were omitted from all calculated performance measures.

Section 2.5.2 explains the structural parameter values maintained for all cases, as well as the parameter values maintained for each of the three tested decision rules RL, FL, and EO. Section 2.5.3 then explains the range of values tested for the treatment factors for each case in Table 2.3.

2.5.2 Maintained Parameter Values

2.5.2.1 Structural parameter values maintained for all cases

As detailed in Section 2.4, the SP Benchmark Model is fully determined, given the utility and production function specifications (2.37) and (2.39) together with the parameter value specifications listed in Table 2.2. These function and parameter specifications are maintained for all cases reported in this study.

As detailed in Section 2.2.4, the constraints and goals of the I structurally-identical consumers in the DM Game depend on the specific settings for $(M^{c,0}_{-1}, u(\cdot), \beta)$. Also, as detailed in Section 2.2.5, the constraints and goals of the J structurally-identical firms in the DM Game depend on the specific settings for $(M^{f,0}_{-1}, q^{stock}_{-1}, F(\cdot), \mu, \kappa^{div})$. All of these functions and parameters have fixed specifications for all cases reported in this study. The utility and production function specifications $u(\cdot)$ and $F(\cdot)$, plus the values of β and q^{stock}_{-1}, are set at the same values as set in Section 2.4 for the SP Benchmark Model, and the values for the remaining parameters are set at the values listed in Table 2.4.

[10]Specifically, these ten seed values were as follows: $\{2012, 2013, 2014, 1, 2, 3, 100, 101, 102, 345\}$.

Table 2.4 Maintained parameter values for the constraints and goals of consumers and firms

Parameter	Value
I	10
J	3
$M_{-1}^{c,0}$	1.00
$M_{-1}^{f,0}$	10.00
μ	0.95
κ^{div}	0.50

The transformation function $\mathbf{TR}_{it}^{c}$ for consumer i in period t postulates that consumer i calculates at time t a reservation wage (2.14) for each current and future period $r \geq t$, which in turn depends on consumer i's expectation for the wage in periods $r \geq t$. Similarly, the transformation function $\mathbf{TR}_{j,t}^{f}$ for firm j in period t postulates that firm j at time t calculates a reservation wage (2.17) and a reservation goods price (2.18) for each current and future period $r \geq t$, which in turn depend on firm j's expectations for the wage and goods price in periods $r \geq t$.

As detailed in Appendix A.2, the methods used by the consumers and firms to form and update these wage and goods price expectations in each period t depend on these agents' prior beliefs regarding wages and goods prices, and also on their memory length, i.e., the number of past periods they take into account when forming these expectations. The prior-belief parameters are set at maintained values, given in Table A.5. However, as will be clarified below in Section 2.5.3, two different settings are tested for the memory length.

2.5.2.2 Parameter values maintained for each decision procedure

The decision domain D^c in (2.13) for each consumer i depends on the grid specifications for Ω and Θ. Also, the decision domain D^f in (2.16) for each firm j depends on the grid specifications for L^f, Γ, Λ, and Ψ. As detailed in Tables A.1 through A.4 in Appendix A.1, two different forms are considered for these grid specifications: namely, a *small* form and a *big* form.

The RL decision rule described in Section 2.3.4 is characterized by a parameter vector $(\mathcal{D}, \mathbf{q}(0), C, e, r)$. The only treatment factor for an RL agent is the recency parameter r; all other parameters are maintained at fixed values.

More precisely, the parameter $\mathcal{D}$ is the cardinality of the decision domain D^c for an RL consumer or D^f for an RL firm. This cardinality is determined by the grid-type specification for D^c or D^f, which is always set to *small* for an RL consumer or RL firm. The vector $\mathbf{q}(0)$ of initial propensities has dimension $\mathcal{D}$. This vector is set equal to a fixed vector $\mathbf{q}^{c,*}$ for an RL consumer and to a fixed vector $\mathbf{q}^{f,*}$ for an RL firm, where these fixed vectors are defined as follows. For an RL consumer, the initial propensity assigned by $\mathbf{q}^{c,*}$ to a decision $d^c = (l^c, \omega, \theta) \in D^c$ is 1.1 if $l^c = 1$ and 1.0 otherwise. For an RL firm, the initial propensity assigned by $\mathbf{q}^{f,*}$ to a decision $d^f = (l^f, \gamma, \lambda, \psi) \in D^f$ is 1.1 if $l^f = l^f_L$ and 1.0 otherwise. Finally, the cooling parameter C is set to 1.0 and the experimentation parameter e is set to 0.95. These maintained values are summarized in Table 2.5.

Table 2.5 Maintained parameter values for RL agents

Parameter	Value
grid-type	small
$\mathbf{q}(0)$	$\mathbf{q}^{c,*}, \mathbf{q}^{f,*}$
C	1.00
e	0.95

The FL decision rule described in Section 2.3.5 is characterized by the vector $\mathbf{Q}_0$ of initial Q-value estimates $\widehat{Q(x,d)}$ as well as by the parameter vector $(\epsilon, \gamma, \alpha)$. The state-space for x is discretized for each FL agent in order to keep computational solution-times manageable. The state $x_{i,t}$ of an FL consumer i at each time $t \geq 0$ is given by his time-t money balance $M^c_{i,t-1}$, discretized into the following three bins: $[0.0, 5.0)$,$[5.0, 10.0)$, $[10.0, \infty)$. The state $x_{j,t}$ of an FL-firm j at each time $t \geq 0$ consists of its time-t money balance M^f_{t-1} and its time-t goods stock q^{stock}_t, each also discretized into three bins, as follows: for the money balance, $[0.0, 50.0)$, $[50.0, 100.0)$, $[100.0, \infty)$; and for the goods stock, $[0.0, 5.0)$,$[5.0, 10.0)$, $[10.0, \infty)$.

The vector $\mathbf{Q}_0$ of initial Q-value estimates is set equal to a fixed vector $\mathbf{Q}^{c,*}$ for an FL consumer and to a fixed vector $\mathbf{Q}^{f,*}$ for an FL firm, where these fixed vectors are defined as follows. For an FL consumer, the initial Q-value estimate assigned by $\mathbf{Q}^{c,*}$ to a state-decision pair (x, d^c), where $d^c = (l^c, \omega, \theta) \in D^c$, is 0.5 if $l^c = 1$ and 0.0 otherwise. For an FL firm, the initial Q-value estimate assigned by $\mathbf{Q}^{f,*}$ to a state-decision pair (x, d^f), where $d^f = (l^f, \gamma, \lambda, \psi) \in D^f$, is 0.5 if $l^f = l^f_L$ and 0.0 otherwise. Finally, the learning parameter γ in (2.27) is set to 0.95 and the greedy parameter ϵ is set to 0.10. These maintained values are summarized in Table 2.6.

Table 2.6 Maintained parameter values for FL agents

Parameter	Value
grid-type	small
$\mathbf{Q}_0$	$\mathbf{Q}_0^{c,*}, \mathbf{Q}_0^{f,*}$
γ	0.95
ϵ	0.10

Implementation details for the EO-ADP and EO-FH decision rules are provided in Appendix A.3. The maintained parameter values for these EO decision rules are also given in Appendix A.3 in order to enable a better understanding of their meaning and role.

2.5.3 Tested Specifications for Case Treatment Factors

As detailed in Appendix A.1, two different settings are tested for the decision-domain grid specifications: namely, a *small* setting and a *big* setting. Although a small grid-type is maintained for both the RL and FL decision procedures, both small and big grid-types are tested for EO agents.

As detailed in Appendix A.2, two different settings are tested for the memory parameter wm used by consumers and firms to adaptively update their expectations. The first setting, $wm = one$, indicates that consumers and firms in each period t only make use of realizations from the previous period $t-1$ to form their expectations for periods $r \geq t$. The second setting, $wm = all$, indicates that consumers and firms in each period $t > 0$ make use of realizations from all previous periods $\{0, \ldots, t-1\}$ to form their expectations for periods $r \geq t$.

Note that all tested cases depend on the setting for wm. This dependence arises because, as detailed in Sections 2.3.2 and 2.3.3, the transformation functions $\mathbf{TR}^c_{i,t}$ and $\mathbf{TR}^c_{j,t}$ mapping consumer and firm period-t decisions into collections of demand and supply functions for periods $r \geq t$ depend on the wage, price, and dividend payment expectations of the consumers and firms, which in turn depend on wm.

For the cases listed along the diagonal in Table 2.3, the tested combinations of values for the treatment-factor parameters are as shown in Tables 2.7 through 2.10. All cross-products of the listed parameter values are tested.

Table 2.7 Tested treatment-factor parameter values for RL agents in cases 1-10

Parameter	Range of Values
r	$\{0.05, 0.10, 0.5, 0.90, 0.95\}$
wm	1, all

Table 2.8 Tested treatment-factor parameter values for FL agents in cases 11-20

Parameter	Range of values
α	$\{0.05, 0.10, 0.50, 0.90, 0.95\}$
wm	1, all

Table 2.9 Tested treatment-factor parameter values for EO-FH agents in cases 23-30

Parameter	Range of values
T	$\{5, 20\}$
wm	1, all
grid-type	small, big

Table 2.10 Tested treatment-factor parameter values for EO-ADP agents in cases 35-38

Parameter	Range of values
wm	1, all
grid-type	small, big

For the remaining cases in Table 2.3, the tested values for the treatment-factor parameter values are as shown in Table 2.11. Superscripts are used to indicate for which decision rule each tested value applies.

Table 2.11 Tested values of treatment-factor parameters for cases 21, 22, 31-34, and 39-44

Parameter	Value
r^{RL}	0.05
wm^{RL}	*all*
α^{FL}	0.05
wm^{FL}	*all*
T^{EO-FH}	20
wm^{EO-FH}	*all*
grid-type^{EO-FH}	*small*
wm^{EO-ADP}	*all*
grid-type^{EO-ADP}	*small*

2.6 Key Simulation Findings for the DM Game

2.6.1 Overview

This section summarizes key DM Game simulation findings for the 44 tested decision-rule cases listed in Table 2.3. Recall that each case in Table 2.3 corresponds to a distinct setting of values for the treatment-factor parameters for that case.

For the most part, we focus attention on utility outcomes for the DM Game consumers since the DM Game firms are merely vehicles to facilitate production. Since different cases involve different planning-horizon lengths, the main ex post performance measure used below for each case k is *average realized single-period utility* $\bar{u}^k$, bounded above and below by two standard deviations $\sigma_{\bar{u}^k}$. Other ex post performance measures used to report results include the *average realized single-period utility for period* t, denoted by $\bar{u}_t^k$, the *average realized cumulative utility through period* t, denoted by $\bar{u}_t^{cumul,k}$, the *average* *realized real wage*, denoted by $\bar{w}^{real,k}$, the *average realized real wage for period* t, denoted by $\bar{w}_t^{real,k}$, and *average realized single-period profits*, denoted by $\bar{\pi}^k$. The calculations for all of these measures are given in Appendix A.5.

Overall, cases with EO-FH agents tend to achieve better performance than cases with only RL, FL, and/or E-ADP agents. However, comparative performance depends strongly on the settings for the treatment-factor parameters. For example, a long memory length covering all previous periods tends to result in better performance than a short (one-period) memory length, all else equal.

We begin this section by focusing on simulation findings obtained for the diagonal cases in Table 2.3, for which the DM consumers and firms all use the same type of decision rule. We then proceed to an examination of the off-diagonal cases in which mixed combinations of decision rules are used.

2.6.2 Findings for the Pure RL Cases 1-10

Consider cases 1-10 in Table 2.3 for which all consumers and firms are RL agents. Each of these cases corresponds to a distinct setting of the RL treatment factors (r, wm) in Table 2.7, taking as given the maintained parameter values in Table 2.5.

As seen in Section 2.3.4, the recency parameter $r \in [0, 1]$ determines the weight $[1-r]$ that is placed on accumulated past single-period utility realizations relative to the weight $[1 - e]$ placed on the most recent single-period utility realization. Since e is set at the maintained value $e = 0.95$, a reduction in r implies an increase in the weight placed on past utility outcomes relative to the weight placed on the most recent utility outcome. A longer memory length wm =*long* should be beneficial for performance in a stationary environment, but it could be harmful to performance in a non-stationary environment. Interestingly, in the DM Game the bulk of the uncertainty faced by each agent is uncertainty regarding the decision-making behavior of other agents. Consequently, the more that the agents settle down in their decision-making selections, the more stationary the environment becomes.

Figure 2.4 reports performance outcomes for cases 1-10 in Table 2.3. The performance of each case k is measured by average realized single-period utility $\bar{u}^k$, and cases are ordered from left to right in ascending performance order.

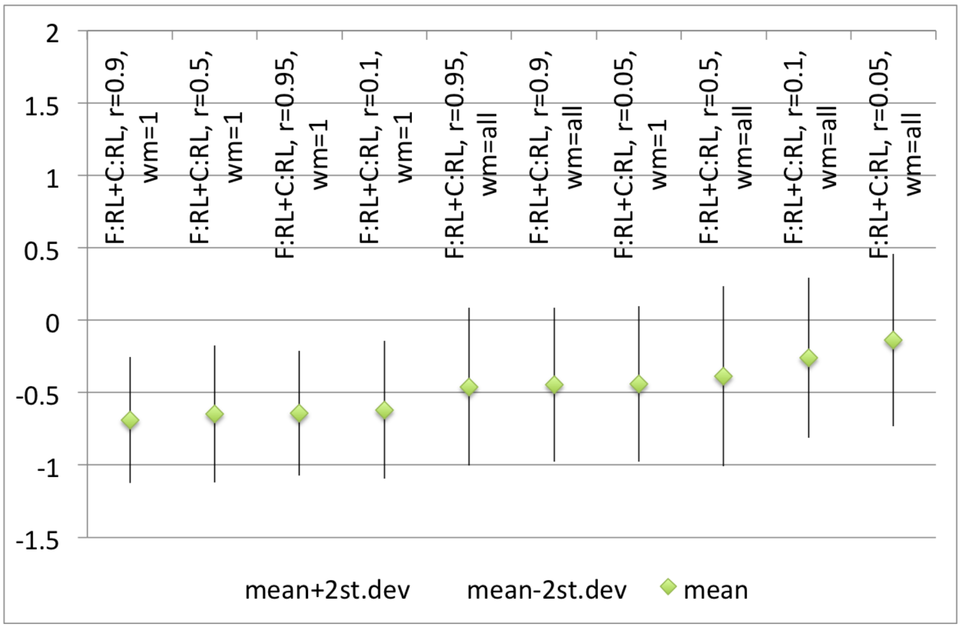

Figure 2.4 Pure RL cases 1-10: average realized single-period utility $\bar{u}^k$ with bounds of $\pm$ two standard deviations $\sigma_{\bar{u}^k}$

Given a longer memory length $wm{=}all$, it is seen that smaller r values (larger weights on past utility outcomes) tend to result in better performance than larger r values. Given a one-period memory length $wm{=}one$, however, a relatively low performance level results for all r values. Moreover, even in the best-performing cases, performance is significantly below 2.08, the stationary per-period utility level (2.43) obtained by the representative consumer in the SP Benchmark Model

2.6.3 Findings for the Pure FL Cases 11-20

Consider, next, cases 11-20 in Table 2.3, for which all consumers and firms are FL agents. Each of these cases corresponds to a distinct setting of the FL treatment factors (α, wm) in Table 2.8, taking as given the maintained parameter values in Table 2.6.

As seen in Section 2.3.5, the update weight $\alpha \in [0, 1]$ determines the weight $[1 - \alpha]$ that is placed on past Q-value estimates relative to the weight α placed on current and anticipated future utility outcomes based on the most recent utility outcome and a new state realization. Since these two weights sum to 1.0, a reduction in α implies an increase in the weight placed on past utility outcomes relative to current and anticipated future utility outcomes.

Figure 2.5 reports performance outcomes for cases 11-20 in Table 2.3. The performance of each case k is measured by average realized single-period utility $\bar{u}^k$, and cases are ordered from left to right in ascending performance order.

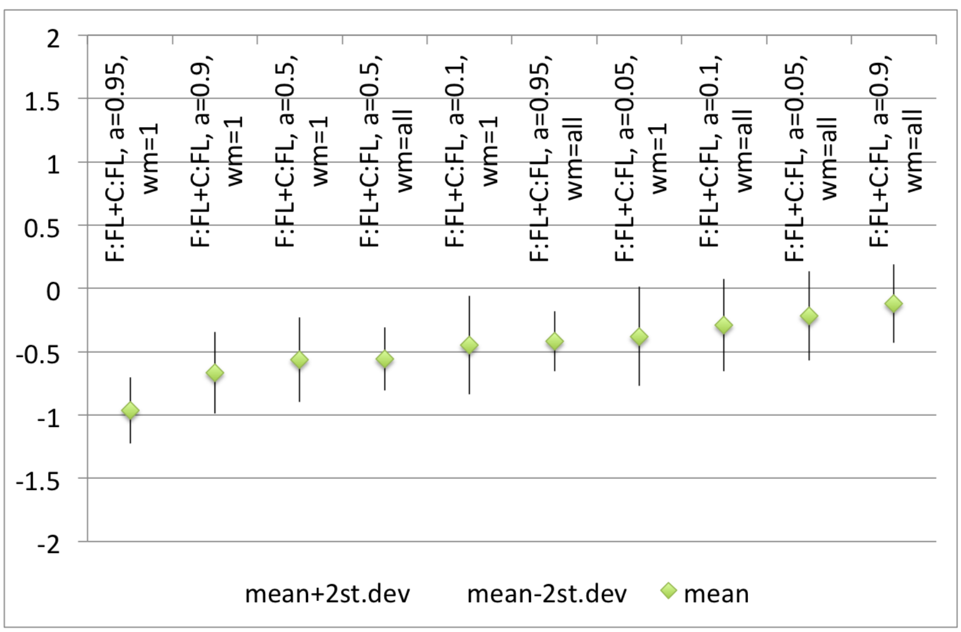

Figure 2.5 Pure FL cases 11-20: average realized single-period utility $\bar{u}^k$ with bounds of $\pm$ two standard deviations $\sigma_{\bar{u}^k}$

Given a longer memory length *wm=all*, it is seen that larger α values (smaller weights on past utility outcomes) tend to result in better performance than smaller α values, although this is not uniformly true. Given a one-period memory length *wm=one*, a relatively low performance level generally obtains regardless of the setting for α, again with exceptions. Indeed, as for the pure-RL cases, even the best-performing pure-FL cases have a performance level that is significantly below 2.08, the stationary per-period utility level (2.43) obtained by the representative consumer in the SP Benchmark Model

2.6.4 Findings for the Pure EO-FH Cases 23-30

Now consider cases 23-30 in Table 2.3, for which all consumers and firms are EO-FH agents. Each of these cases corresponds to a distinct setting of the EO-FH treatment factors T, wm, and grid-type in Table 2.9, taking as given the maintained parameter value NDrawsFH=10 discussed in Appendix A.3.2.

A longer forecasting horizon T means that the EO-FH agent is more anticipatory. This could be beneficial if the agent's anticipations are an accurate reflection of future uncertainties, but it could be harmful if not. Restricting the number of potential decision selections by specifying grid-type=*small* rather than grid-type=*big* increases the sampling density, i.e., the frequency with which each potential decision is tried. On the other hand, grid-type=*small* results in a cruder approximation of the decision domain, which could prevent the EO-FH agents from determining their truly best decisions.

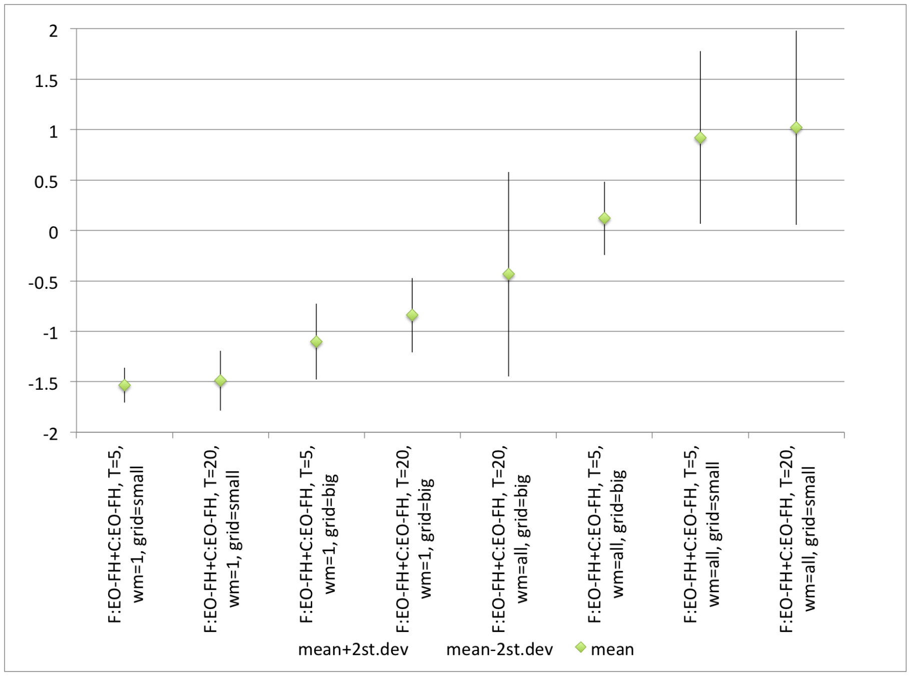

Figure 2.6 Pure EO-FH cases 23-30: average realized single-period utility $\bar{u}^k$ with bounds of $\pm$ two standard deviations $\sigma_{\bar{u}^k}$

Figure 2.6 reports performance outcomes for cases 23-30 in Table 2.3. The performance of each case k is measured by average realized single-period utility $\bar{u}^k$, and cases are ordered from left to right in ascending performance order.

Given a one-period memory length wm=one, performance is relatively low regardless of the grid-type or the length T of the forecasting horizon. However, given a longer memory length wm=all, it is seen that having a small grid-type results in better performance than a large grid-type.

Moreover, for wm=all and grid-type=$small$, the longer forecasting horizon T=20 yields slightly better performance than the short forecasting horizon T=5. Indeed, as indicated by the standard deviation bounds in Fig. 2.6, for this combination of treatment factors the average realized single-period utility level $\bar{u}^k_t$ attained in some periods t comes close to matching the stationary single-period utility level 2.08 achieved by the representative consumer in the SP Benchmark Model. This occurs despite the rather simplistic

Monte Carlo method used by EO-FH agents to handle their uncertainty regarding future wages, prices and dividends.

Given the relatively good performance of the EO-FH decision procedure under some treatment-factor specifications, it is interesting to delve deeper into the underlying dynamics. Time-series for utility and real wage outcomes are depicted below for two illustrative cases: (i) a "good" case 26 with T=20, wm=all, and grid-style=$small$; and (ii) a "bad" case 29 with $T = 20$, wm=one, and grid-style=big.

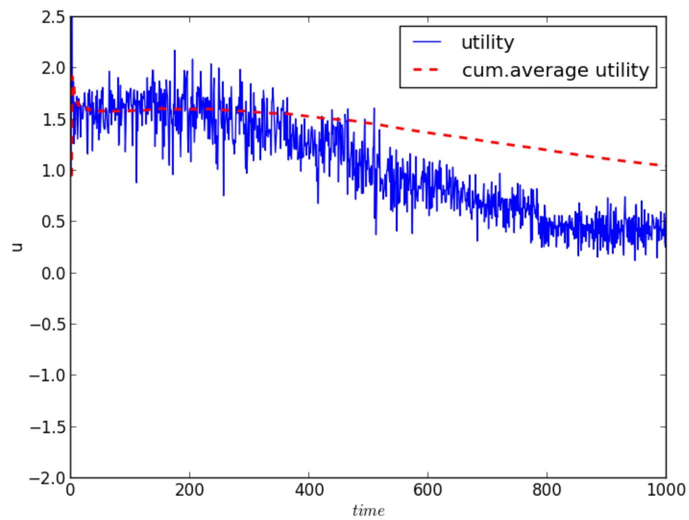

Figure 2.7 Pure EO-FH case 26: average realized single-period utility $\bar{u}_t^{26}$ for period t and average realized cumulative utility $\bar{u}_t^{cumul,26}$ through period t, over successive periods t

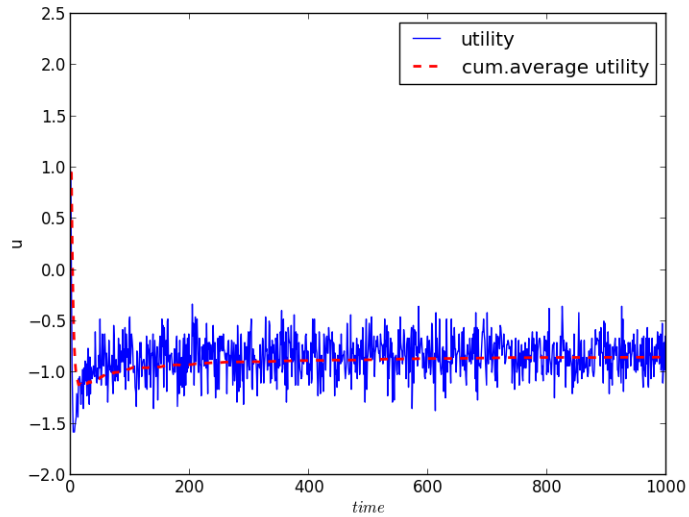

Figure 2.8 Pure EO-FH case 29: average realized single-period utility $\bar{u}_t^{29}$ for period t and average realized cumulative utility $\bar{u}_t^{cumul,29}$ through period t, over successive periods t

For the "good" case 26, depicted in Fig. 2.7, the average realized single-period utility $\bar{u}_t^{26}$ eventually stabilizes at a level of about 0.5. For the "bad" case 29, depicted Fig. 2.8, the average realized single-period utility $\bar{u}_t^{29}$ quickly stabilizes at a much lower level of about -1.0.

The behavior of the real wage reflects overall macroeconomic performance. For the "good" case 26, it is seen in Fig. 2.9 that the average realized real wage $\bar{w}_t^{real,26}$ appears to be stabilizing at a level of about 0.30. In contrast, for the "bad" case 29, it is seen in Fig. 2.10 that the average realized real wage $\bar{w}_t^{real,29}$ rapidly drops towards zero.

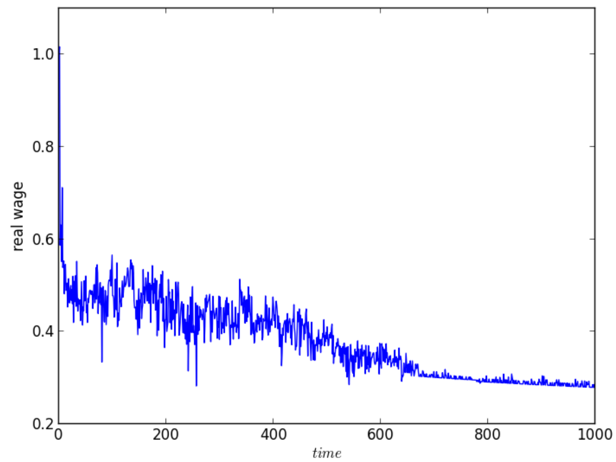

Figure 2.9 Pure EO-FH case 26: average realized real wage $\bar{w}_t^{real,26}$ for period t, over successive periods t

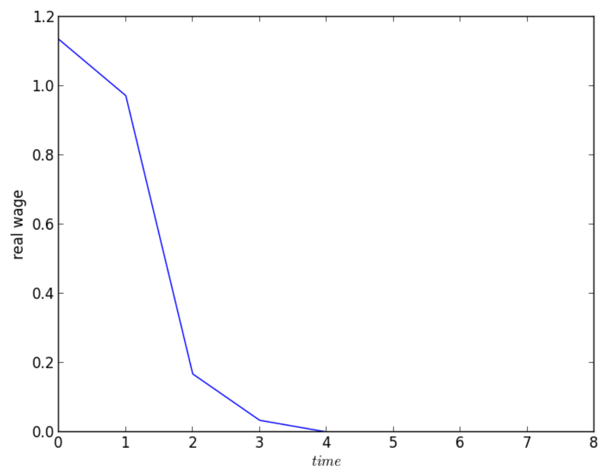

Figure 2.10 Pure EO-FH case 29: average realized real wage $\bar{w}_t^{real,29}$ for period t, over successive periods t

2.6.5 Findings for the Pure EO-ADP Cases 35-38

Consider cases 35-38 in Table 2.3, for which all consumers and firms are EO-ADP agents. Each of these cases corresponds to a distinct setting of the EO-ADP treatment factors wm and grid-type in Table 2.10, taking as given the maintained parameter values listed in Table A.6.

Figure 2.11 reports performance outcomes for these four cases. The performance of each case k is measured by average realized single-period utility $\bar{u}^k$, and cases are ordered from left to right in ascending performance order.

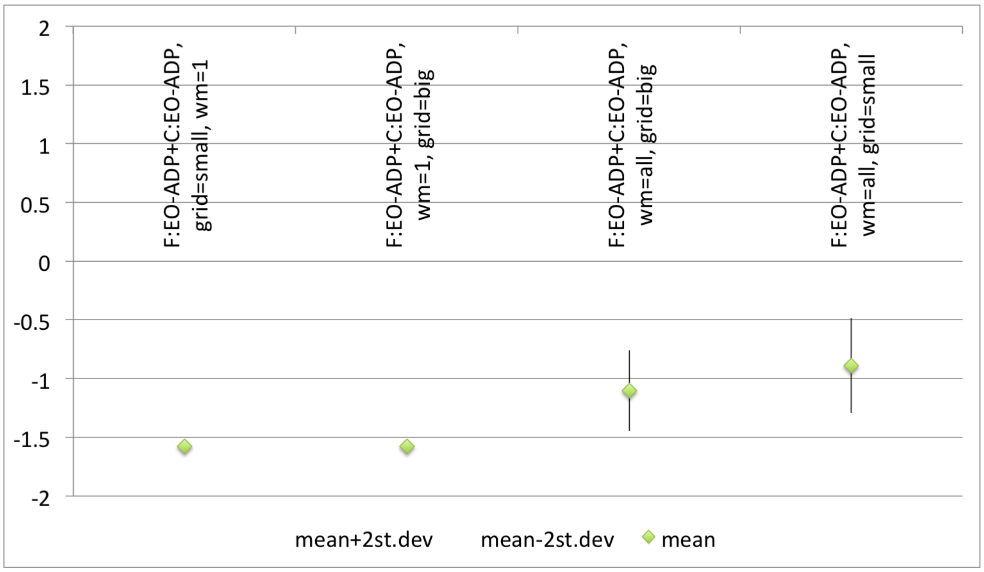

Figure 2.11 Pure EO-ADP cases 35-38: average realized single-period utility $\bar{u}^k$ with bounds of $\pm$ two standard deviations $\sigma_{\bar{u}^k}$

EO-ADP performance is clearly better with a longer memory wm=*all* than with a one-period memory wm=*one*. Moreover, given a longer memory, performance is slightly better with grid-style=*big* in comparison with grid-style=*small*. Overall, however, a low performance level is attained for all tested settings of the EO-ADP treatment factors in comparison with the overall performance attained using the RL, FL, and EO-FH decision procedures.

2.6.6 Findings for Mixed Combinations of Decision Rules

From a social welfare point of view, it is only consumer utility outcomes that matter in the DM Game. However, the players in the DM Game are utility-seeking consumers and profit-seeking firms, where the latter act on behalf of their shareholders (who receive their profits as dividend payments) but not consciously on behalf of consumer welfare per se.

It is therefore of interest to construct consumer and firm payoff matrices for the DM Game, interpreting the alternative possible decision procedures RL, FL, EO-FH, and EO-ADP as possible pure strategy choices for these players.

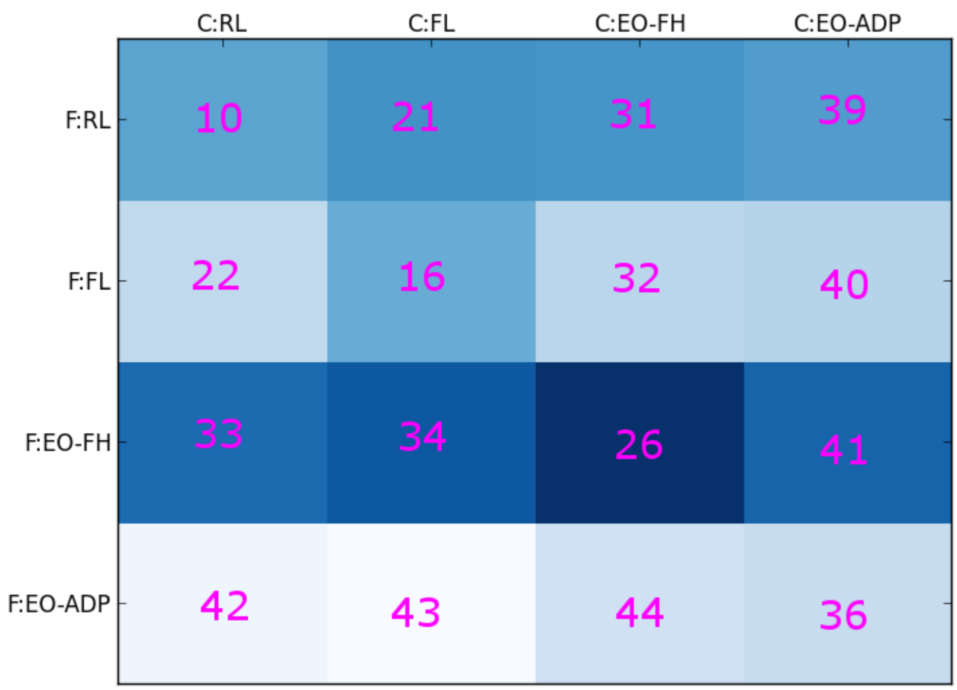

Figure 2.12 Consumer payoff matrix for the DM Game reporting average realized single-period utility $\bar{u}^k$ for the indicated cases k. A darker shade of color indicates a higher value for $\bar{u}^k$.

We therefore tested the off-diagonal cases in Table 2.3 representing mixed combinations of decision procedures. We then used the performance outcomes obtained for these off-diagonal cases together with the performance outcomes obtained for the diagonal cases to construct DM-Game payoff matrices, one for consumers and one for firms, under the restriction that all consumers use the same decision procedure and all firms use the same decision procedure.

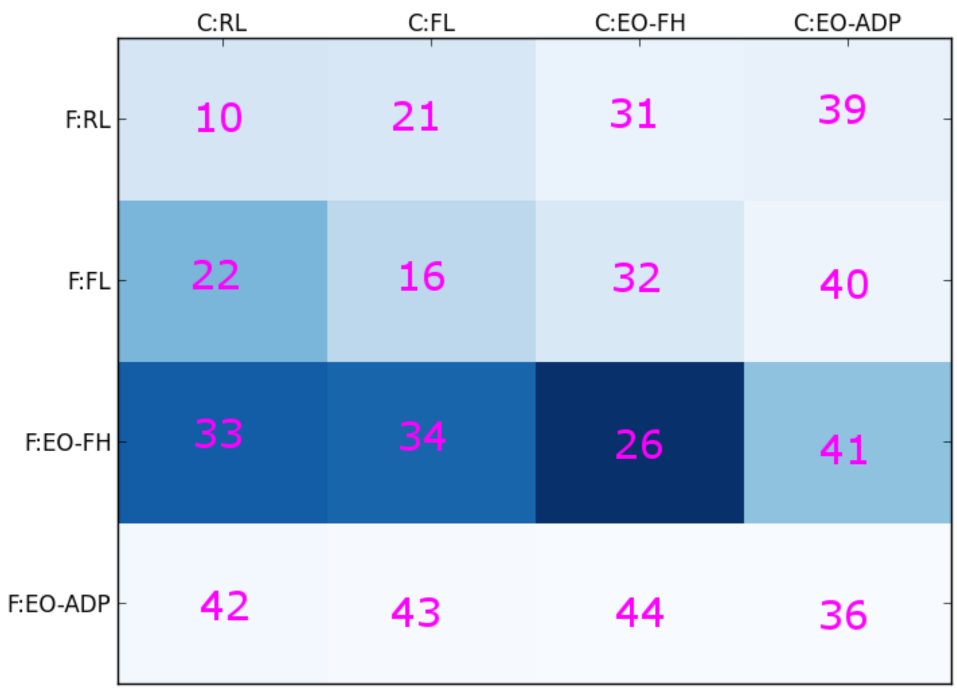

Figure 2.13 Firm payoff matrix for the DM Game reporting average realized single-period profits $\bar{\pi}^k$ for the indicated cases k. A darker shade of color indicates a higher value for $\bar{\pi}^k$.

The consumer payoff matrix, depicted in Fig. 2.12, reports the average realized single-period utility $\bar{u}^k$ attained by consumers for each indicated case k, with darker shades of color corresponding to higher values of $\bar{u}^k$. The firm payoff matrix, depicted in Fig. 2.13, reports the average realized single-period profits $\bar{\pi}^k$ attained by firms for each indicated case k, with darker shades of color corresponding to higher values of $\bar{\pi}^k$.

It is important to note the following non-standard aspect of these payoff matrices. For each pairing of consumer and firm decision procedures along the diagonals, the treatment-factor parameters are selected in an attempt to permit each agent type to do as well as possible in this pairing. This is reflected in the fact that, in contrast to Table 2.3, only single cases are considered along the diagonals.

As seen from the firm payoff matrix in Fig. 2.13, EO-FH is a dominant strategy for firms, given the particular case selections and treatment-factor specifications used to form this payoff matrix. Interestingly, as seen from the consumer payoff matrix in Fig. 2.12, this is not true for consumers. For example, the best response of consumers to a firm choice of FL is to choose FL, not EO-FH. Nevertheless, it is also seen from these two payoff matrices that (EO-FH, EO-FH) is a Pareto optimal Nash equilibrium

2.7 Conclusion

This study explores the comparative performance of constructively rational decision-making procedures in the context of an otherwise standard macroeconomic model with intertemporally optimizing consumers and firms. These decision-making procedures range from simple reactive reinforcement learning to sophisticated adaptive dynamic programming (ADP) techniques.

A key finding is that the best macroeconomic performance tends to result for cases in which agents use the EO-FH procedure and have long memories. The EO-FH procedure determines approximate intertemporal utility and profit solutions by means of direct search, using a finite rolling planning horizon. In particular, EO-FH with long memory tends to dominate the tested RL procedure based on Roth-Erev reactive reinforcement learning, the tested FL procedure based on Q-learning, and the tested EO-

ADP procedure based on an adaptive dynamic programming method for value function approximation.

However, to date, only a small number of parameter values have been explored for each of these decision-making procedures, and there is no guarantee that the best parameter settings for the DM Game environment have been used. Moreover, further testing is needed to clarify the effects of memory length, forecasting horizon, and grid-point density specifications for decision domains, and the interactions among these specifications, in alternative economic environments.

Clearly, then, much further study is needed to understand the ramifications of requiring consumers and firms in macroeconomic models to be constructively rational, in accordance with their real-world counterparts. In particular, a large gap exists between *constructive rationality*, i.e., basing decisions on one's own beliefs, information, and attributes, and *constructive optimality*, i.e., the assurance that the combination of decision rules in use by agents satisfy some stated optimality property, such as Pareto optimality.

Nevertheless, a primary goal of this study has already been accomplished: namely, to provide a proof-of-concept demonstration that consumers and firms in computational models can be implemented as forward-looking learners and intertemporal planners whose decision-making results in sustained economic activity, despite the absence of top-down coordination devices such as rational expectations and global market clearing conditions.

Another important goal accomplished by this study is the development of a modular, extensible, and scalable macroeconomic framework that facilitates the comparative analysis of different institutional structures populated by a mix of agents with diverse decision-making procedures. In subsequent work, the range of considered structures and procedures will be extended to permit consideration of more realistic features, such as the inclusion of a central bank and a commercial banking system.

CHAPTER 3. ECONOMIC SURVIVAL AS A FUNCTION OF BEHAVIORAL RULES AND INFORMATION PREFERENCES

Agent-based computational economics (ACE) is a diverse set of approaches and methods that could be used to study a range of problems and analyze consequences of behavior under conditions that could not be solved analytically. One of such questions is optimal behavior in a changing environment, when the amount of information and the learning opportunities are severely restricted. In this Chapter, I studied a range of combinations of learning and choice policies available to an agent, ranging from simple rules to more sophisticated approaches based on expected utility maximization with expectations in the form of a Bayesian network. I have found out that a three-level Bayesian network coupled with approximate optimization techniques might perform on a par with the exact solution and correct belief specifications.

3.1 Introduction

One of the established approaches to model uncertainty and choices under uncertainty is to assume Von Neumann–Morgenstern utility function and to solve the resulting optimization problem using an expected utility of an agent. Such an approach, however, assumes deep knowledge about the world that people occupy, or, at least, about the main characteristics of this world. This assumption is hardly a realistic one. It would be more reasonable to assume that people might be perceiving the world they are acting in as one of many possible worlds with the corresponding priors on the distribution of such possible worlds. This approach is researched in the paradigm of ambiguity preferences.

The question of preferences under ambiguity, and the corresponding behavioral consequences has recently become an active research area. Risk aversion (or uncertainty aversion) is a standard part of an economic model, but incorporation of ambiguity aversion is very limited. A recent review of models with preferences over ambiguous outcomes is given in Epstein and Schneider (2010). They review dynamic models of ambiguity-risk aversion and also show that the time-consistent dynamic preferences in a form of RU (recursive utility) has corresponding static preferences. They also discuss a limited number of applications for these models. The main models they review are Recursive SEU, which corresponds to static Subjective Expected Utility (SEU), Recursive Multiple-Priors with Maxmin Expected Utility (MEU) as a static preferences, and Recursive Smooth Ambiguity Model with corresponding static preferences given in Klibanoff et al. (2005). A generalized version of RSU can be found in Hayashi and Miao (2010).

However, these utility function representations have not been tested in a laboratory environment. Only recently have static representations been tested in Ahn et al. (2007), where they found some evidence that the tendency to equate demands for securities that pay off in the ambiguous states could be more easily accommodated by the α-MEU (α-Maxmin Expected Utility) model than by the SEU model.

Besides experimental evidences that people behave differently under uncertainty and ambiguity, it was also shown in Hsu et al. (2005) that decision making in uncertain and ambiguous environments activate different parts of the brain. Neural activity while taking ambiguous decisions was also investigated in Bach et al. (2009). A number of experiments studied heterogeneity in ambiguous preferences. Borghans et al. (2009) showed that men and women have different ambiguity preferences. Keck et al. (2010) studied group decisions making in an ambiguous setup. As in the case of uncertainty, it was shown that framing matters for ambiguous choices in Ho et al. (2002). Maffioletti et al. (2009), and Trautmann et al. (2009) showed that the preference reversals effect is valid for an ambiguous preferences.

Testing of savings/consumption decisions under uncertainty and a general survey of macro experiments can be found in (Duffy, 2008). Carbone and Hey (2004) showed that subjects are generally unable to solve the dynamic optimization problem, but might react in a correct direction to changes in the environment. Hey and Knoll (2011) conducted experiments to define decision rules used to solve savings/consumption by subjects.

This work is trying to bind together, on the one hand, experimental results that show that agents use simple rules to make savings/consumption choices, and, on the other hand, a highly advanced mathematical model that tries to explain agent choices. In this model, agents are allowed to have better specified beliefs as compared to the simple ambiguous beliefs. They also implement different possible simplification to the optimization problem.

3.2 Structure of the Model

3.2.1 The Main Question

The model was developed to test a range of hypotheses that deal with optimal choices under uncertainty. When making intertemporal choices, people try to find a balance between the best possible behavior and the uncertainty that surrounds the results of their choices. If we knew the exact rules that govern the economy, we could, arguably, choose the best possible actions (barring the issue of game interactions that will complicate such choices). But what happens if we do not know much about the world we live in and have to learn about it along the way? What will be the best belief structure we could assume, and how should we learn about the world? Is this belief structure universal, or does it depend on the particulars of the world? Those questions are too broad to be answered in a single paper. Many researches tried to offer partial solutions to them. A lot of effort was devoted to defining and researching the implications of different utility functions that might be useful for ambiguous situations. Another branch of research deals with researching learning under uncertainty. The model proposed in this work combines these issues and tries to answer the big question of the best possible behavior under incomplete information and limited learning opportunities.

The model introduces Bayesian networks for the belief structure and approximate optimization algorithms for making choices under uncertainty. Both of these tools help to define beliefs for a broader range of situations and serve as a vehicle for a performance evaluation of different approximation techniques that people are using or should be using when trying to survive in a stochastic world. The implemented approximate optimization algorithm is scalable and could be used in other applications, especially when full scale optimization algorithms are infeasible.

The first part of the model deals with the belief representation. The Bayesian updating of beliefs is used because it is the only instrument that is consistent from the statistical point of view. The general belief structure is also formalized through the Bayesian network. This generalization allows for a simultaneous specification of the different assumptions on agent beliefs. One-level network corresponds to simple beliefs and Von Neumann-Morgenstern utility. Two-level network describes ambiguous preferences. Three and more levels correspond to a higher ()relative to an ambiguous beliefs) order of beliefs.

The second part of the model is designed to test different optimization algorithms. As a benchmark, the complete search on a grid is used. It is highly time consuming, not scalable, but provides exact solutions. As an alternative, an approximate dynamic programming algorithm adopted from Powell (2011) is tested.

Other possible variables are controlled in the following way. Multiagent learning interactions are excluded by allowing only one agent to make decisions at one moment in time. Decision feedback loops are excluded by subjecting an agent to an exogenously defined stochastic process that is independent of his decisions. Information is limited by excluding sampling and providing an agent only with a historical realization of the stochastic process on returns.

To sum it up, there is one agent that tries to optimize the expected utility by making a consumption/savings choice. This agent observes returns on savings that are realizations of a stochastic process. To allow for an explicit solution to the optimization problem, each agent lives for only 3 periods and after that is replaced by an identical agent. This new agent may inherit some or all of the properties of the previous agent.

3.2.2 Structure of the Worlds

The underlying driving force of all agent decisions is a stochastic process that defines returns on savings. For simplicity, this process is assumed to be discrete. The set of possible returns includes two returns $\{r_1, r_2\}$ with the corresponding distribution $[p_{w_i,r_1}, p_{w_i,r_2}]'$.

As there are only two possible realization of returns, $p_{w_i,r_2} = 1 - p_{w_i,r_1}$. Here w_i is a subscript that denotes the world. Each world is characterized by a probability distribution over returns. An agent may find himself living in a fixed world, i.e. with a simple fixed probability distribution over returns, or may live in an environment where a probability distribution is itself subject to change. The later option allows for an inclusion of a deterministic or stochastic switching processes for returns.

The agent knows the form of a probability distribution and the exact set of possible returns, but has to learn probabilities for these returns. The only information that the agent has are period-by-period realized returns on his savings. Since in general the agent does not know which stochastic process is generating returns, he has to form some believes about the possible probability of getting one of the returns. This uncertainty is usually captured by ambiguous beliefs and an appropriate utility function. In terms of Bayesian networks, it will correspond to a two-level Bayesian network. A one-level

Bayesian network will describe an agent that thinks that he knows the exact distribution that generated the returns. A three-level Bayesian network will describe an agent that thinks that he has no knowledge about possible probabilities at all.

To simplify the analysis and at the same time make it more illustrative, the specific forms of belief structures are introduced. An agent with ambiguous beliefs with 50% probability believes that he lives in the correct world, and with 50% in the world where probabilities are reversed. This specific form allows for an easy introduction of the three-node Bayesian network.

The one-node network will include cases where an agent believes with 100% probability that he is in the right or the wrong world. Those specifications will be used for the testing purposes, as they represent an absolutely correct and an absolutely incorrect prior, respectively. The composition and the structure of an agent are kept intentionally simple to allow for the focus on beliefs and choice algorithms.

3.2.3 Population and Inheritance

To keep computations feasible, it is assumed that at each moment in time only one agent is active (for a total of T periods) and is replaced by a new agent at the end of a lifespan. A new agent receives an endowment at the beginning of his life, and forms his beliefs. These beliefs could be his own, and thus independent of others, or inherited from the old generation. The form of a belief inheritance is an another treatment factor in the simulations. In general, an agent could also inherit wealth from an older generation, but as there is no stimulus for an agents to care about the younger generation, nothing will be left to pass to them. The specific forms of preferences and beliefs are described below.

3.3 Agent Preferences and Updating of Beliefs

3.3.1 Agent Preferences

Each agent is assumed to live for $T = 3$ periods, and therefore his utility at time t can be represented in the following form:

$$U_{t,T,T^{rem}} = \sum_{\tau=t}^{t+T^{rem}} E_{\tau}\beta^{\tau}u\left(c_{\tau}\right), T^{rem} = 3, 2, 1 \tag{3.1}$$

As the agent becomes older, his utility function shrinks to exclude the past period. The period utility function $u\left(c_t\right)$ has a CES form:

$$u\left(c_t\right) = \frac{c_t^{1-\theta_u}}{1-\theta_u}, \theta_u \in [0.1, 4] \tag{3.2}$$

Each agent has an endowment w_0 that it has to distribute over his life-time (T). There is a random interest rate for the savings. The distribution of returns for this interest rate is the source of uncertainty for an agent.

At each moment t agent faces budget constraints that correspond to the remaining periods of his life:

$$M_{t+\tau} \leq M_{t+\tau-1}\left(1 + r_{t+\tau}\right) - c_{t+\tau}, \tau = 0, \cdots, T^{rem} \tag{3.3}$$

Here the money holding in the new period $M_{t+\tau}$ are the money holding in the previous period plus the interest income for this money holdings with an interest rate $r_{t\ +\ \tau}$ minus the consumption in this period $c_{t\ +\ \tau}$. The initial endowment defines the amount of money at the beginning of the agent's life:

$$M_{t-1} = w_0 \tag{3.4}$$

if agent begins his life at time t.

Only consumption and saving decisions (no borrowing) are allowed, and there is no additional endowment after the initial period for each agent (when he is young).

There is also a non-negativity constraint on consumption $c_t \geq 0$.

3.3.2 Expectations

The agent forms the expectations about possible realizations for an interest rate based on the information he gets in each period. The only new information he can get is the past realized interest rate, because no sampling of interest rates is allowed as described in Section 3.2.1. Thus, at each period his information set ΔI_t consists of only one point, namely the realized interest rate.

$$\Delta I_t = \{r_t\} \tag{3.5}$$

To improve the quality of his choices, the agent also forms prior beliefs about the possible world structure. These beliefs are updated in a consistent way using Bayesian updating. This assumption is a rather demanding one, as it could be computationally consuming, to the degree when people have to use heuristics to cope with such levels of complexity. Such heuristics have been found in the experiments, but we do not consider them in the current study. Instead, we concentrate on the benefits and the disadvantages of a statistically consistent update of beliefs, not heuristics.

A number of different setups for beliefs structure is studied. The simplest possible structure arises when the agent has the information about the exact world he is living in (and believes this information to be true). The only source of uncertainty in this case is the interest rate. Formally , the agent believes that he knows true w_i, and thus knows p.d.f. for the returns, i.e. P_{t,r,w_i} for each moment t. In this case, he can use the true return distribution P_{t,r,w_i} in his utility estimation.

$$U_{t,T,T^{rem}} = \sum_{\tau=t}^{t+T^{rem}} E_{P_{\tau,r,w_i}} \beta^{\tau} u\left(c_{\tau}\right), \; T^{rem} = 3,2,1 \tag{3.6}$$

Another scenario arises if the agent realizes that the information he has about the true world he is living in may be incorrect. In this case, he has to form beliefs over the possible worlds he may be in. These beliefs take the form of a Bayesian prior P_w - p.d.f. for possible probabilities of the different worlds that the agent might be acting in.

A general form of the utility function in this case is

$$U_{t,T,T^{rem}} = V\left(\{P_{t+\tau,w}\}_{\tau=1}^{T^{rem}}, \{P_{t+\tau,r,w_i}\}_{\tau=1}^{T^{rem}}, \{c_{t+\tau}\}_{\tau=1}^{T^{rem}}\right) \tag{3.7}$$

where V is some utility function that takes as parameters the consumption stream $\{c_{t+\tau}\}_{\tau=1}^{T^{rem}}$ and the beliefs $\{P_{t+\tau,w}\}_{\tau=1}^{T^{rem}}$ over the return distributions $\{P_{t+\tau,r,w_i}\}_{\tau=1}^{T^{rem}}$.

The specific form of V assumed in this paper is the following;

$$U_{t,T,T^{rem}} = E_{P_w} \sum_{\tau=t}^{t+T^{rem}} E_{\tau,P_{t+\tau,r,w_i}} \beta^{\tau} u\left(c_t\right),\ T^{rem} = 3,2,1 \tag{3.8}$$

In the simple case of this model, we assume that the agent has a two point discrete prior for worlds one and two that are characterized by the reversed probabilities. For the case of ambiguous beliefs, we assume that at the beginning of a simulation the agent believes that both of the world structures are of equal probability. These beliefs correspond to a two-level Bayesian network. The diagram below illustrates his beliefs.

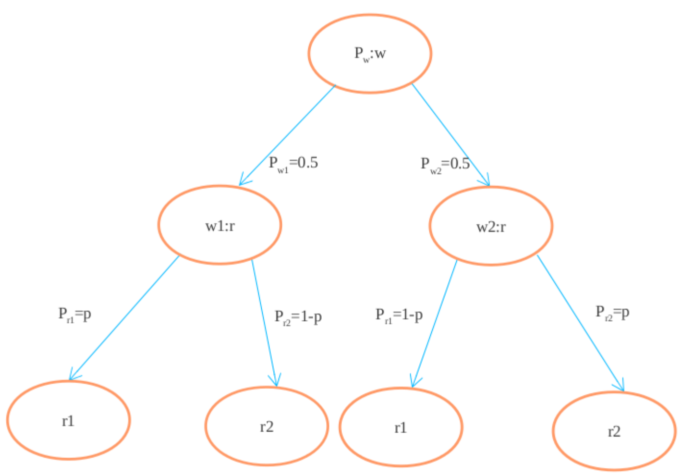

Figure 3.1 Structure of ambiguous beliefs

At each period some new information is received, and the prior for the world structure is updated, thus producing new estimates for p_{w1} and, respectively, for $p_{w2} = 1 - p_{w1}$. This updating is done using the Bayesian approach in the following way:

$$p_{t+1,w_1} = \frac{P\left(W_{t+1} = r_i | w = w_1\right) p_{t,w_1}}{P\left(W_{t+1} = r_i | w = w_1\right) p_{t,w_1} + P\left(W_{t+1} = r_i | w = w_2\right)\left(1 - p_{t,w_1}\right)} \tag{3.9}$$

After that, the updated expectations $E_{P_w}\left(x | I_t\right)$ are used in equation (3.8).

The last tested specification is a three-level Bayesian network as shown below:

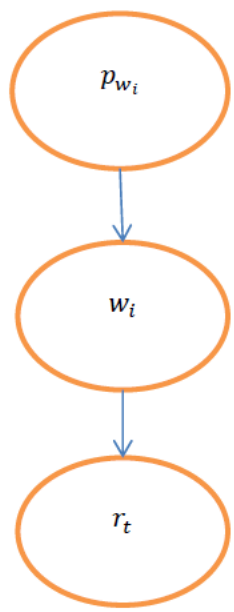

Figure 3.2 Three-node Bayesian network of beliefs

This form of the belief specification adds another level of uncertainty, in this case on the P_w prior. Instead of assuming that it is 50/50 distribution, he believes that p_{w1} itself is uniformly distributed over the interval $[0, 1]$. Here $p_{w1} = \theta$ is the probability that the first world structure is true, and $F_\theta \sim uniform[0, 1]$ is the distribution for this probability. For the simulation purposes, it is later discretized into $\{\alpha_n\}_{n=1}^{N}$, where $P(\theta = \alpha_n) = \frac{1}{N}$. $N = 10$ is hold fixed for all simulations.

When some new information arrives, the distribution parameters are updated as following

$$
\begin{aligned}
P(\alpha | r = r_k) &= \frac{P(\alpha, r = r_k)}{P(r = r_k)} \\
&= \frac{\sum_{\theta, r = r_k} P(\alpha) P(\theta|\alpha) P(r = r_{tk}|\theta)}{\sum_{\theta} P(r = r_k|\theta) \sum_{\alpha} P(\theta|\alpha) P(\alpha)}
\end{aligned}
\tag{3.10}
$$

This hierarchical Bayesian prior describes the situation when the agent faces not only uncertainty and ambiguity, but also a true unknown situation and is aware of that. The expected utility for this case is modified to include expectations over all levels of priors.

Given his preferences and beliefs, the agent tries to maximize the expected utility over his remaining lifetime, and spreads the initial endowment or, later in life, the money holdings in the best possible way. We tested different optimization algorithms in combinations with different beliefs structures to assess which ones perform better in terms of the average utility. Two algorithms used in the simulations were the complete search and the optimization using Approximate Dynamic Programming.

3.4 Algorithms for Optimization

3.4.1 Approximate Dynamic Programming Algorithm

Approximate Dynamic Programming Algorithm (called ADP from now on) was adopted from Powell (2011). Before describing the specific realization of an algorithm, we describe the model in more general terms. Let X_t be the state at time t. In our case it is

$$X_t = (M_t, \mathcal{F}_t, H_t) \tag{3.11}$$

where M_t is the amount of money on hand, $\mathcal{F}_t$ are beliefs about the structure of the world (for the case of ambiguous beliefs, it is p_{w1}, and for the case of the three-level Bayesian prior it is $\{\alpha_n\}_{n=1}^{N}$) and H_t is a hidden state. H_t is used for storing the information about the specific realization of a "world structure" random variable from the two/three-level Bayesian network and is required for accurate Monte-Carlo simulations.

This state formulation is itself a simplification. A more general state specification would include all past realizations of the interest rate as a part of history, so the state would be $X_t = (M_t, \mathcal{F}_t, H_t, I_t)$. Here it is assumed that all the information I_t is encompassed in beliefs $\mathcal{F}_t$.

Denote by d_t the decisions of the agent at time t. In this model, it is the share of the income that the agent decides to consume. Given the choice d_t, the consumption c_t of the agent equals

$$c_t = d_t M_t \tag{3.12}$$

In this model, d_t was discretized in the interval $[0, 1]$ with the number of discretization points $CS\_N\_discret$ equal to 10.

$W_{t+1}(\omega)$ is the realization of random variables. In our model, $\omega \in \{r_1, r_2\}$ is a realization of random returns.

The rule for updating the state can in general be expressed in the following way:

$$X_{t+1} = TR\left(X_t, d_t, W_{t+1}(\omega)\right) \tag{3.13}$$

Given the state X_t, the decisions d_t and the realization of random variables $W_{t+1}(\omega)$, the new state X_{t+1} is decided using TR mapping. In the current model, TR includes rules for updating belief and resources. In the case of ambiguous beliefs, equation (3.9) is used. In the case of the three-level priors, equation (3.10) is used.

Resources include the money on hand, which are updated according to (3.3) and (3.12).

Besides the state and rules for updating the state, another general part of the model is a period contribution. In this model, it is the period utility:

$$C\left(X_t, d_t\right) = u\left(c_t\right) = u\left(d_t M_t\right) \tag{3.14}$$

The decision takes the form of the share of money to be spent. This form is a simplified linearization of a more general decision rule, which should be of the form $d_t\left(X_t\right) = D^{\pi}\left(X_t\right)$, where the decision depends on the full state, which includes the beliefs.

Approximate Dynamic Programming Algorithm is trying to estimate the value function that is defined in the following way:

$$V_t\left(X_t\right) = \arg\max_{d_t \in D_t}\left(C\left(X_t, d_t\right) + \beta\,\mathbb{E}_t\left(V_{t+1}\left(TR\left(X_t, d_t\right)\right)\right)\right) \tag{3.15}$$

In this model, the basis functions approximation for value function is used. The decision is chosen such that

$$d_t = \arg\max_{d_t \in D_t}\left(C\left(X_t, d_t\right) + \gamma\,\mathbb{E}\left(\sum_f \theta^{\pi}_{tf}\phi_f\left(TR\left(X_t, d_t\right)\right)\right)\right) \tag{3.16}$$

where D_t is the set of possible decisions.

In this linearization, θ^{π}_{tf} are specific to some set of policies π, the coefficients in a linearization of the value function. In our model, the agent lives over T periods and makes decisions every period, therefore, T value functions are needed. This means that the linearization parameters need to be indexed by time. $\phi_f, f = 1, ..f_N$, are basis functions, with f_N being total number of them. The simplest case uses linear basis functions. This simplification was implemented in this study, with $N = 1$, and $\theta^{\pi,0}$ given in the Table B.1.

With these simplifications, the value function approximation is given by

$$V_t\left(X_t\right) = \theta_{0,V} + \theta'_V X_t \tag{3.17}$$

Because of the simple structure of the state given in equation (3.11), it was possible to further simplify the linearization with

$$\theta'_V = [\theta_t, 0, 0] \tag{3.18}$$

and

$$\theta_{0,V} = 0 \tag{3.19}$$

Given these simplifications, only the current amount of money on hand is taken into account, but not the beliefs about the possible world structures and, of course, the hidden state.

For the sub-step of the search for an optimal policy, given the value function estimation, the coarse-grained complete search is used. Other algorithms could not provide the necessary accuracy of the estimation.

The exact implementation of the algorithm is described in Appendix B.2.

3.5 Testing Schemes and Results

3.5.1 Testing Schemes

The main dimensions for testing are the behavioral rules, represented by optimization algorithms, and the belief structure. All other simulation parameters were chosen to better illustrate the performance of the different behavioral rules and were kept fixed during simulations.

The simulation environment is mostly defined by the returns structure. The following returns structure (where p_{r_i} is the probability of getting the return equal to r_i) was tested.

Table 3.1 Tested probability-returns combinations

$\frac{Description}{Code}$		r_1	p_{r_1}	r_2	p_{r_2}
P-R2	w_1	3.0	p	-0.3	1-p
	w_2	3.0	1-p	-0.3	p

This combination of returns was chosen because it is an efficient representation of the high risk environment. Probabilities p_{r_i} were tested in the range of $\{0.1, 0.2, ..., 0.9\}$

As for the belief structures, the following specifications were tested. In the case of ambiguity preferences, it was the 50/50 split for the prior and the completely correct specification. The corresponding probabilities p_{w_i} are given below.

Table 3.2 Tested specifications for the ambiguity beliefs

Code name	p_{w_1}	p_{w_2}
A1	0.5	0.5
A2(correct prior)	1	0

In the case of the three-level Bayesian network, the uniform third level prior was tested.

The preference parameters were fixed at the following level:

Table 3.3 Tested parameters for the preferences

Parameter	**Value**
θ_u	3.0

These parameter values represent a strong form of risk aversion for the CES period utility function.

The other used parameters were the inheritance of beliefs and value function estimations, the time preferences, the initial endowment, the lifespan of an agent, and some technical parameters for the ADP and complete grid search algorithm. The number of simulation runs J and the corresponding random generator seeds were fixed. The specification of these parameters is given in Appendix B in Table B.1.

The following combinations of the belief structure and the optimization algorithm were tested:

Table 3.4 Tested combinations of beliefs and decision algorithms

Name	**Description**
A,CS	ambiguous beliefs, complete search
A,ADP	ambiguous beliefs, ADP solution algorithm
UK,CS	three-level prior, complete search
UK,ADP	three-level prior, ADP solution algorithm

3.5.2 Results

Below are presented the heatmaps that comparatively describe the performance of the different strategies when p, the probability of getting r_1, is changing. The performance is measured by the average realized utility $\bar{u}$ over the length of a simulation $LRun$. The initial learning period with the length of $LOmit = 30$ is excluded.

The average realized utility of an agent is calculated in the following way:

$$\bar{u} = \left[\sum_{j=1}^{NSeeds} \sum_{t=LOmit}^{LRun} u_t \right] / \left[NSeeds \cdot (LRun - 29) \right] \tag{3.20}$$

Below are given the comparative results in the form of the heatmap for all the tested parameter combinations:

Figure 3.3 Heatmap for the belief-algorithm simulation results, all tested cases (darker is better)

The results for the cases where agents implement the same complete grid search algorithm for the decision making, but differ in their belief structures, are presented below:

Figure 3.4 Heatmap for belief-algorithm simulation results, cases with the complete search algorithm
(darker is better)

The results for the cases where agents implement the same approximate algorithm for decision makin,g but differ in their belief structures, are presented below:

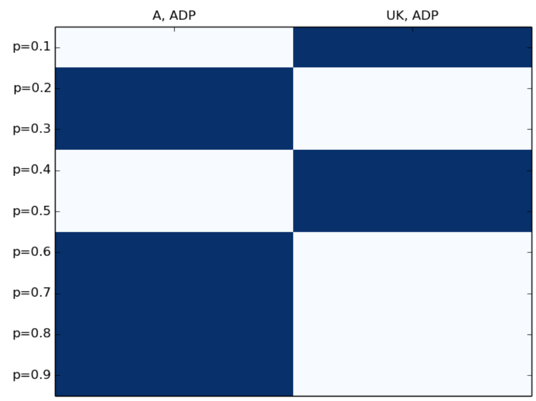

Figure 3.5 Heatmap for the belief-algorithm simulation results, cases with the approximate search algorithm
(darker is better)

One can see from the heatmaps that having a more advanced type of preferences (acknowledging the unknown) helps to improve the overall performance in some, but not all, cases.

Surprisingly, it pays out to be less smart in most cases, if the approximate algorithm is used. On the other hand, the performance is split when the precise algorithm is used. A more advanced belief structure is beneficial for an agent when the environment itself is beneficial. On the contrary, it pays out to be more conservative, if the environment is not so beneficial to an agent.

The approximate algorithm performs worse than the complete search, as could be expected, but the performance differences decrease as the environment becomes less extreme. It needs to be mentioned that the approximate algorithm used here was implemented in a straightforward form. Nevertheless, it still delivered a comparable performance for some combinations of the initial parameters.

As a future direction of the research, it would be interesting to see if the share of income that is set aside for savings stabilizes in such simulations. If it is, then it may indicate that fixed rules might have evolved as a suitable strategy for making choices in an unstable environment.

3.6 Conclusions

In this chapter, it has been shown that more accurate beliefs are not always beneficial to the agent, and sometimes being overly optimistic proves to be more beneficial when the environment is conductive to such biases. Given that each environment favors a specific combination of beliefs and a specific optimization algorithm, it can be expected that over time the strategies best fitted for the corresponding environment survive. This could explain why in reality we observe seemingly fixed behavioral rules. These rules might be the surviving ones in a changing environment. It is also informative to see that an approximate algorithm, even in its crudest form, can perform reasonably well in a moderately risky environment.

There are few potential improvements that could be made over the tested algorithms and the general specification of the problem. It may be interesting to try to develop a more generalized approach that would allow introduction of more than two alternative world specifications in the case of the agent with a tree-level Bayesian belief structure. Another possible way of developing this simple model would be to introduce advanced linearization schemes for the value function approximation. The performance of this algorithm heavily depends on the choice of linearization functions and other parameters.

Overall, this simple model proved to be a useful tool for the analysis of a sophisticated belief structures and their effects on the agent's performance in a world initially unknown to him.

CHAPTER 4. IS QUANTITATIVE EASING ENOUGH?

In this chapter, we develop an agent-based model that includes the banking sector and use this model to analyze effects of non-conventional monetary policy used by central banks. This model includes main sectors typical of a developed economy and has separate central bank, government and banks agents. This extension of the model in comparison to the models discussed in chapters 1 and 2 allows us to analyze the effects of the central bank intervention on short- and long-term liquidity markets in detail. We have found out that the institutional structure and regulations, as well as expectation formation rules, dominate over whatever policies the central bank implements. From the practical viewpoint, given the current goal of returning economies to their potential long-term growth rates, it is clear that complex institutional reforms need to be implemented. Relying on a very limited toolset of central banks is not enough to achieve this goal.

4.1 Introduction

In the wake of 2008 financial crisis, central banks around the world implemented different policies to try and get economies back to long-term growth rates. One of these policies was an expansion of balance sheets and targeting interest rates other than short-term ones. These efforts took different forms. For example, the Federal Reserve implemented “credit easing” by buying mortgage-backed securities and “operation twist” that changed the term structure of balance sheets. Other banks, such as the European Central Bank, expanded balance sheets by using long-term repo operations with the collateral being mostly bank loans. The Bank of Japan has been buying government securities for a long time until now. The Bank of England bought government bonds from the non-banking sector.

For the purposes of this paper, we define quantitative easing policies (QE) as buying the government bonds with the goal of influencing interest rates and the economic activity. It is only one of the possible choices of the definition, and the rationale behind it was the need to have an empirical point of reference. With this definition, the U.S. economy could be used as an example economy for the model, as this definition of QE corresponds to what the Federal Reserve is currently implementing.

A general overview of different QE and conventional policies can be found in Joyce et al. (2012). These policies are an increasing part of more general efforts to formulate a macroprudential set of policies, an overview of which can be found in Galati and Moessner (2013).

As for the formal studies of such policies, only the DSGE models have been used to analyze them so far. For example, Curdia and Woodford (2011) extends the standard New Keynesian model to include the central bank and its balance sheet to analyze the effects of unconventional versus interest rate policies. Gertler and Karadi (2013) continues this work and introduces a generalized approach to modeling unconventional policies. Other models were introduced in Bernanke and Reinhart (2004), Curdia and Woodford (2010), Gertler and Karadi (2011), Christiano (2010). Another approach is to extend a big-scale DSGE model to include the financial sector and use it to make policy estimates, as in De Resende et al. (2013). All these models suffer from the deficiencies intrinsic to the DSGE approach. They could not properly represent heterogeneous iterations between agents subject to network effects, do not model explicitly institutional structures of an economy, and also do not take into account limited information available to agents, especially when considering that QE policies are new and not well understood by participating agents. Also, the banking sector is not properly described in DSGE models, as the money and credit creation process is not modeled explicitly, as well as the corresponding macroprudential and banking regulations.

On the other hand, the current ABM model have not been used to analyze details of the banking sector and their interactions so far. There is some research into modeling banking and financial sector with ABM models, such as Ashraf et al. (2011), or Dosi et al. (2013), but none of these models were designed to include forward-looking agents that try to optimize their goals subject to constraints.

Our model presented in this chapter was developed to address these issues, both the lack of forward-looking behavior in ABM models and insufficient institutional and informational modeling in the DSGE models. The structure of the model was designed to mimic the U.S. economy and the FED policies, but at the same time was significantly simplified to keep the complexity at a manageable level. The description of the model is given below.

4.2 General Description of a Model

4.2.1 Introduction to the Model

This model was designed to serve as a tool to study out-of-equilibrium short and medium term effects of a range of policies (quantity and others) pursued by central banks. The main focus is on careful outlining and depiction of institutional conditions for the banking sector operations and decisions.

Let us start by describing the main agents and decisions in the model, and move on to highlighting interesting trade-offs faced by the agents and the modeler.

In this model, banks take the main stage, while other agents in the model are designed to better highlight banks' decisions. As a source of institutional definitions, the U.S. economy was chosen. The computational efficiency considerations required some simplifications on the part of the decision procedures for some of the agents. In addition to that, certain institutions were simplified or eliminated from the model. The resulting set of agents and institutions is presented below.

4.2.2 Introduction to Agents in the Model

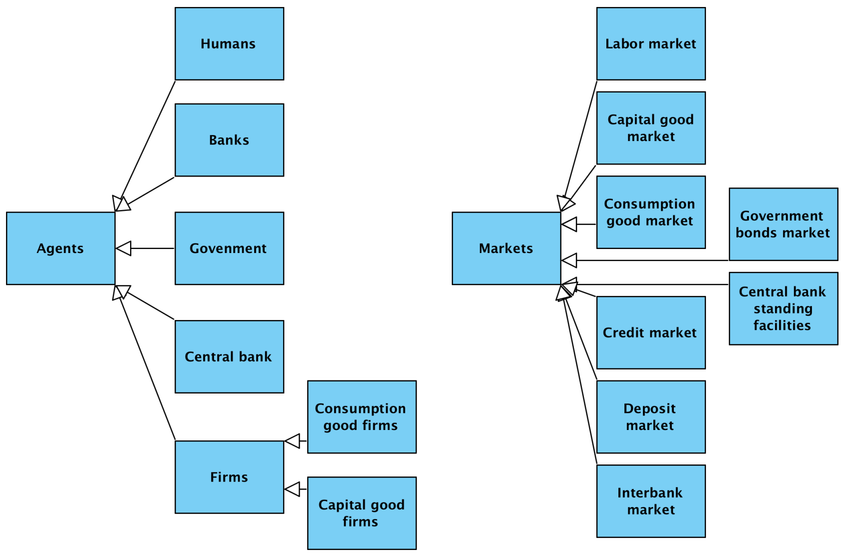

Figure 4.1 Agents and markets in the model

In this economy there are humans, consumption good producing firms, capital goods producing firms, banks, the central bank and the government. Details on particular choices and life cycles of these agents are presented further in the text.

Markets in this model were limited to the labor, capital, consumption good, credit, deposit, interbank and government bonds markets. There is only one market for the government bonds. This limitation required imposing additional assumptions on the form of the QE policies.

There are standing facilities, provided by the central bank and available to all banks. Supporting institutions include the payment and legal systems.

All these agents interact in a particular way on the markets. The sequence of interactions is presented below.

4.2.3 Main Events

The events that take place in the economy and their time ordering are presented below.

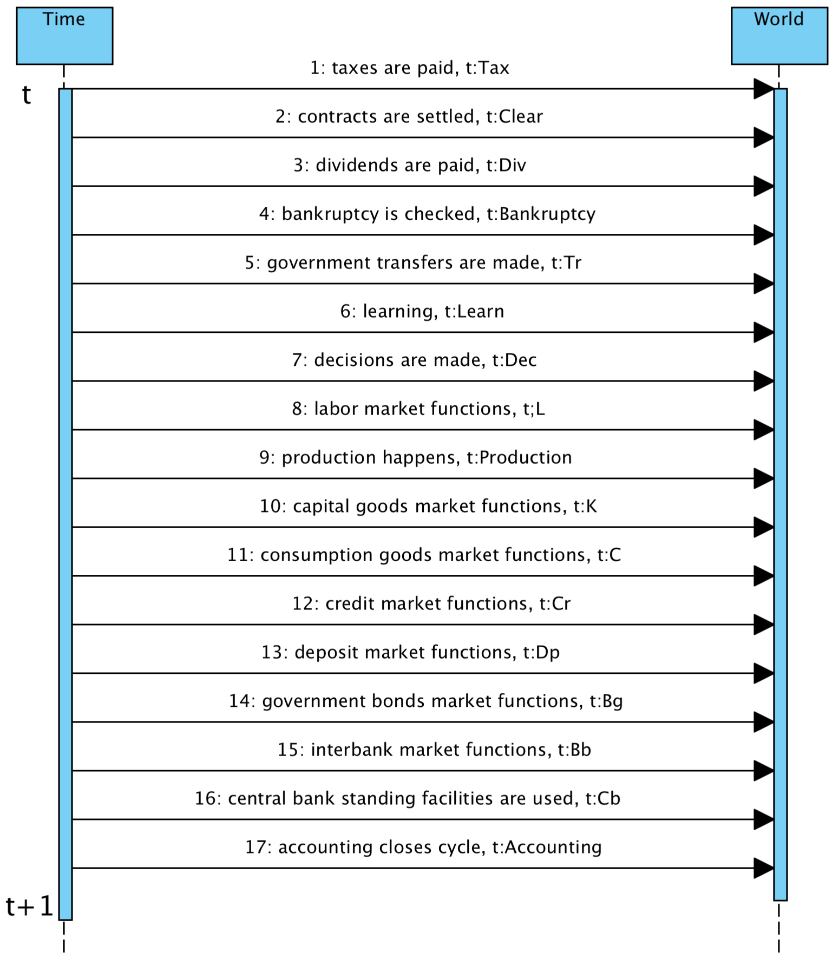

Figure 4.2 Sequence of events in the model

At the beginning of period t, taxes to the government by banks, firms and humans are paid. Next, labor, credit, deposit and government bonds contracts are cleared, and the necessary payments are made. After that, the dividends by firms and banks are paid in equal shares to all the humans in the economy. Checks for bankruptcy as described in Section 4.3.5.11 are performed next. After that, the agents receive the information about the previous results of market clearing in period $t-1$, and they update their expectations. After updating the expectations, new decisions are made. After the labor market functions and contracts are signed, production takes place and all other markets function in a sequence. First, the capital goods market; then, the consumption goods market, the credit, deposit, government bonds markets and the interbank market. After that, the standing facilities of the central bank are used, if required. Finally, accounting of the period and final calculation of the period profits and utilities are done.

The labor and government bonds markets function with the periodicity specified by the parameters $f\_Hk = 5$ and $f\_Bg = 6$. Taxes and dividends are paid with the frequencies $f\_taxes = 3$ and $f\_div = 5$.

4.3 Agents and Markets

4.3.1 Banks

In a typical economy of a developed country, there are multiple types of financial intermediaries and a wide rage of financial instruments. The scope of this model was constrained to a subset of these agents and instruments. The criterion for inclusion and aggregation was to preserve accurate functioning of the money markets and all of the related markets as much as possible.

Banks are the sole financial intermediaries in this model. Correspondingly, only few of many types of different financial instruments are chosen to be representative tools for facilitating financial intermediation in this simplified model. All derivative instruments are out of the scope of the model, as well as the active portfolio management.

Banks are the core agents that act on all the markets in the economy: the credit, deposit, interbank markets, the labor and goods markets, the market for the central bank funds, and the bonds market.

The typical sequence of events is described below. Banks pay all the due money to the other agents in the economy. After that, their solvency may be checked. The information is collected, and the expectation formation rules are updated according to the new information. Banks make their choices, and after that all the markets function in the same order as described in Section 4.2.3.

4.3.1.1 Bank constraints and goals

When banks make their choices for the coming period, the goal of bank v at time t is to maximize the expected profit in the form:

$$\max_{d \in D^b} E_{v,t} \sum_{r=t}^{\infty} \mu^{r-t} \left[\Pi_t \left(\mathbf{X}_r, d, \mathbf{W}_r \right) \right] \tag{4.1}$$

subject to no-bankruptcy conditions

$$M_{cb,t:Account} \geq 0 \tag{4.2}$$

and the resource accumulation constraints given in equations (4.5) and (4.4).

Here $\mathbf{X}_t = (x_r)_{r=0}^{t}$ is the generalized state of the bank at time t, with x_t being the bank's state of assets and liabilities at time t. $\mathbf{W}_t = (w_r)_{r=0}^{t}$ is the history of all the realizations of the random variable for the bank, which includes clearing prices on the goods $p_{c,t:C}$ and the capital market $p_{k,t:K}$, wages $w_{t:L}$, interests $i_{c,t:Cr}$ on the credit market and $i_{d,t:Dp}$ on the deposit market, the interest on the interbank market $i_{bb,t:Bb}$, the price on the government bonds market $p_{bg,t:Bg}$ and the amount of payments ps_t that goes through the bank over the course of period t.

The period profit can be represented in the following way:

$$\begin{aligned}\Pi_t = i_{c,t}\left(c_{t:Cr}, i_{c,t}, \mathbf{c}_{t-1}\right) \\ -i_{d,t}\left(d_{t:Dp}, i_{d,t}, \mathbf{dp}_{t-1}\right) \\ +i_{bb,t}\left(bb_{t:Bb}, i_{bb,t}\right) \\ -i_{cb,t}\left(cb_{t:Cb}, i_{cb,t}\right) \\ +i_{bg,t}\left(bg_{t:Bg}, i_{bg,t}\right) \\ -w_t\left(l_t, \mathbf{w}_{t:L}, \mathbf{hk}_{t-1}\right) \\ -dep_t\left(k_t, \mathbf{p}_{K,t:K}, \mathbf{k}_t\right)\end{aligned} \tag{4.3}$$

In equation (4.4), $i_{c,t}$ is the current period income from credit operations. This includes the interest income that is acquired in the current period from extended credits c_t and is generally equal to $c_t \cdot i_{c,t}$, and includes the interest payments from loans extended in the previous periods that are still outstanding. Here $\mathbf{c}_{t-1}$ includes the information about all the loans extended in periods r, where $0 \leq r < t$.

$i_{d,t}$ is the current period expenses on the accepted deposits. This includes the interest expenses that are acquired in the current period from the accepted deposits d_t, and which are generally equal to $d_t \cdot i_{d,t}$, and also includes the interest payments on the deposits accepted in the previous periods that are still outstanding. Here $\mathbf{d}_{t-1}$ includes information about all the deposits accepted in periods r, where $0 \leq r < t$.

$i_{bb,t}$ is the current period interest payments on the interbank loans. This includes the interest payments that are acquired in the current period from the interbank activity bb_t, and which are generally equal to $bb_t \cdot i_{bb,t}$. Also let $\mathbf{bb}_{t-1}$ include the information about all the loans extended or accepted in periods r, where $0 \leq r < t$.

$i_{bcb,t}$ is the current period interest payments on the central bank loans. This includes the interest payments that are acquired in the current period from using the standing facilities cb_t, and which are generally equal to $cb_t \cdot i_{cb,t}$. Also let $\mathbf{cb}_{t-1}$ include the information about all the loans extended or accepted in periods r, where $0 \leq r < t$.

$i_{bg,t}$ is the current period income from the government bond holdings. This includes the interest expenses that are due in the current period from the bonds on hand bg_t, and is generally equal to $bg_t \cdot i_{bg,t}$. Also let $\mathbf{bg}_{t-1}$ include the information about the government bonds market operations in periods r, where $0 \leq r < t$.

w_t is the current period expenses on labor. They are defined by the currently employed amount of labor l_t and the promised wages $\mathbf{w}_{t:L}$. Here $\mathbf{hk}_{t-1}$ includes the information about all the labor contracts signed in previous periods r, with $0 \leq r < t$, that are still active.

dep_t is the current period expenses on capital. They are defined by the current depreciation of capital, which is a function of the current capital stock k_t and the cost of acquiring capital $\mathbf{p}_{k,t:K}$. Here $\mathbf{k}_{t-1}$ includes the information about all the capital purchases in previous periods r, with $0 \leq r < t$.

There are two major parts in the period profit. The first part is the income(or loss) from banking operations such as providing credit, accepting deposits or trading on bonds market. The second part sums the costs associated with being able to act on the markets. The specific formulations for the profit calculations are given in the code.

The production function $F_v\left(l_t, k_t\right)$, where l_t is employed labor and k_t is capital, generally equal to $K_{t:0}$, requires banks to have at least $L^b_{min} = b\_F\_F\_min\left[0\right]$ of labor and $K^b_{min} = b\_F\_F\_min\left[1\right]$ of capital to be able to act on the credit markets. The parameters $b\_F\_F\_min\left[0\right]$ and $b\_F\_F\_min\left[1\right]$ are given in Appendix C in Table C.1. At the same time, the payment system actions and the deposit market interactions could be carried out without satisfying minimum production requirements.

The capital accumulation equation, given the stock of capital at the beginning of the period $K_{t:0}$ and purchases of capital goods $q_{K,t:K}$, is defined as follows:

$$K_{t+1:0} = \left(1 \;-\; \delta_{dep}\right)\left(K_{t:0} + q_{K,t:K}\right) \tag{4.4}$$

Bank v at the beginning of period t also has money balances at the central bank $M_{cb,t:0}$. These money balances change during period t under the influence of the payment orders submitted by the other agents to the bank. Depending on whether the payee and receiver of the payment have accounts at the same bank or at different banks, or either of them is central bank or the government, the payments are processed differently. After netting all the payments the resulting ps_t and the use of the central bank standing facilities $cb_{t:Cb}$ define the end-of-period money balances $M_{v,cb,t:Accounting} = M_{v,cb,t+1:0}$.

$$M_{v,cb,t+1:0} = M_{v,cb,t:0} + ps_{v,t} + cb_{v,t:Cb} \tag{4.5}$$

Banks also keep track of all the signed contracts, such as labor $\mathbf{hk}_t$, credit $\mathbf{c}_t$ etc., all of which are defined and described above.

Also, at each period bank v calculates its capital $Capital_{v,t}$. Each bank has a fixed stock value $Stock_v$. It also pays dividends $div_{v,t:Div}$ and taxes $tax_{v,t:Tax}$. Given these payments, the capital is calculated as following:

$$Capital_{v,t} = Stock_v + \sum_{r=0}^{t} \Pi_r(\cdot) - \sum_{r=0}^{t} div_{v,r:Div} - \sum_{r=0}^{t} tax_{v,r:Tax} \tag{4.6}$$

Every period the bank also calculates the value of $Assets_t$ in the following way:

$$Assets_{t:r} = M_{cb,t} + bg_{t:r} \cdot p_{Bg,t-1:Bg} + c_{c,t}(c_{t:r}, \mathbf{c}_{t-1}) + k_{t:r} \cdot p_{K,t-1:K} \tag{4.7}$$

where $bg_{t:r} \cdot p_{Bg,t-1:Bg}$ is current valuation of the bond holdings, $c_{c,t}(c_{t:r}, \mathbf{c}_{t-1})$ is the amount of outstanding credits, $k_{t:r} \cdot p_{K,t-1:K}$ is the valuation of the current capital stock, $M_{cb,t}$ are money balances held at the central bank.

All the relevant variables, such as credits, amounts of goods, etc., are subject to non-negativity constraints.

4.3.1.2 Decision domain and transformation functions for banks

To achieve the goal of maximizing the expected profit, the bank makes choice d from the decision domain D^b described below.

$$D^b = \Theta^{Cr} \otimes \Omega^{Cr} \otimes \Omega^{Dp} \otimes \Theta^{Bg} \otimes \Omega^{Bb} \tag{4.8}$$

where:

- the elements of $\Theta^{Cr} = \{\theta_1^{Cr}, \ldots, \theta_{Cr}^{Cr}\}$ satisfy $0 \leq \theta_1^{Cr} < \ldots < \theta_{Cr}^{Cr}$
- the elements of $\Omega^{Cr} = \{\omega_1^{Cr}, \ldots, \omega_{Cr}^{Cr}\}$ satisfy $0 < \omega_1^{Cr} < \ldots < \omega_{Cr}^{Cr}$
- the elements of $\Omega^{Dp} = \{\omega_1^{Dp}, \ldots, \omega_{Dp}^{Dp}\}$ satisfy $0 < \omega_1^{Dp} < \ldots < \omega_{Dp}^{Dp}$
- the elements of $\Theta^{Bg} = \{\theta_1^{Bg}, \ldots, \theta_{Bg}^{Bg}\}$ satisfy $0 \leq \theta_1^{Bg} < \ldots < \theta_{Bg}^{Bg}$
- the elements of $\Omega^{Bb} = \{\omega_1^{Bb}, \ldots, \omega_{Bb}^{Bb}\}$ satisfy $0 < \omega_1^{Dp} < \ldots < \omega_{Dp}^{Dp}$

The labor market decision is defined by the production function. The bid submitted to the labor market is formed in the following way. The bid price $w_{L,t:L}^{bid}$ is

$$w_{L,t:L}^{bid} \;=\; 2 \cdot w_{t-1}^{e} \tag{4.9}$$

and the bid quantity $q_{L,t:L}^{bid}$ is

$$q_{L,t:L}^{bid} \;=\; b\_F\_F\_min\,[0] \tag{4.10}$$

The capital market decision is defined by the production function. If the current amount of capital $K_{t:0}$ is lower than the amount required by the production function $b\_F\_F\_min\,[1]$ given in an Appendix C in Table C.1, then the bid is submitted to the capital market. The bid is formed as in Section 2.15 with the target amount of money to spend equal to $M^{bid,K}$, defined as

$$\ulcorner b\_F\_F\_min\,[1] \;-\; K_{t:0} \urcorner \cdot p_{K,t-1}^{e} \;+\; 0.5 \cdot \sigma_{t-1}^{2,e,K} \tag{4.11}$$

The credit market decision includes the choice of the interest rate gap to the central bank rate ω^{Cr} and the amount of credits to be extended (as a share of assets θ^{Cr}). The bank forms its ask in the following way. The interest rate is:

$$i_{Cr,t:Cr}^{ask} \;=\; i_{cb,t-1:Cb} \;+\; \omega^{Cr} \tag{4.12}$$

where $i_{cb,t-1:Cb}$ is the central bank interest rate on standing facilities. The maximum credit available to agents is $c^{ask}_{t:Cr} = \theta^{Cr} \cdot Assets_{t:Cr-1}$. Given the submitted interest rates by all the banks and the counterparties (borrowers), the equilibrium rate is defined and the market clearing is attempted. During this attempt, each creditor needs to check the actual amount of the extended credit given target debt/assets ratio for agents. In the case of firms, it is the ratio of credit/assets, in the case of humans it is the ratio payments/past average income. Details of this check are described in Section 4.3.5.4. The length of a credit contract c_length_Cr is fixed in simulations for all agents and time periods.

The deposit market decision includes a gap to the central bank rate on the deposit market. The bank submits the following bid price:

$$i^{bid}_{Dp,t:Dp} = i_{cb,t-1:Cb} + \omega^{Dp} \tag{4.13}$$

and the bid quantity

$$d^{bid}_{Dp,t:Dp} = 0.1 \cdot Assets_{t:Dp-1} \tag{4.14}$$

After all the bids and asks are submitted to the deposit market clearing house, the equilibrium rate is calculated, and deposits are made. The length of a deposit contract c_length_dp is fixed, and there is no option to break the contract.

Another market to consider is the market for government bonds. In the current version of the model, there is only one type of bonds that are traded by banks and the central bank. There are primary and secondary markets for government bonds. All banks are allowed to participate on both markets. For the government bonds market, the choices are the share of the assets to have as government bonds θ^{Bg}. Given this choice, the bank decides if it wants to buy or sell government bonds. It buys bonds if the current amount of bonds $bg_{t:0}$ is lower than the desired amount

$$\frac{\theta^{Bg} \cdot Assets_{t:0}}{p_{Bg,t-1:Bg}} \tag{4.15}$$

The bid includes the bid price $p^{bid}_{Bg,t:Bg-1}$ defined in the following way:

$$p^{bid}_{Bg,t:Bg} \;=\; p^{e}_{Bg,t-1} \tag{4.16}$$

and the quantity to buy $q^{bid}_{Bg,t:Bg-1}$:

$$q^{bid}_{Bg,t:Bg} \;=\; bg_{t:0} \;-\; \frac{\theta^{Bg} \cdot Assets_{t:0}}{p_{Bg,t-1:Bg}} \tag{4.17}$$

If the difference between $bg_{t:0}$ and (4.15) is positive, then the bank submits an ask to the market with the ask price $p^{ask}_{Bg,t:Bg}$:

$$p^{ask}_{Bg,t:Bg} \;=\; p^{e}_{Bg,t-1} \tag{4.18}$$

and the quantity to sell $q^{bid}_{Bg,t:Bg-1}$:

$$q^{ask}_{Bg,t:Bg} \;=\; \frac{\theta^{Bg} \cdot Assets_{t:0}}{p_{Bg,t-1:Bg}} \;-\; bg_{t:0} \tag{4.19}$$

Finally, for the interbank market the decisions include the gap to the announced central bank rate, and it is assumed that all the available money at the accounts at the central bank are offered on the interbank market. If the accounts are negative, then banks first try to borrow the necessary sums on the interbank market. So if $M_{cb,t:Bb-1} \geq 0$, the bank submits an ask to the interbank market, with the price $i^{ask}_{Bb,t:Bb}$:

$$i^{ask}_{Bb,t:Bb} \;=\; i_{cb,t-1:Cb} \;+\; \omega^{Bb} \tag{4.20}$$

and the quantity $bb^{ask}_{Bb,t:Bb}$:

$$bb^{ask}_{Bb,t:Bb} \;=\; M_{cb,t:Bb-1} \tag{4.21}$$

If $M_{cb,t:Bb-1} \leq 0$, the bank submits a bid to the interbank market, with the price $i^{bid}_{Bb,t:Bb}$:

$$i^{bid}_{Bb,t:Bb} \;=\; i_{cb,t-1:Cb} \;+\; \omega^{Bb} \tag{4.22}$$

and the quantity $bb^{bid}_{Bb,t:Bb}$:

$$bb^{bid}_{Bb,t:Bb} \;=\; M_{cb,t:Bb-1} \tag{4.23}$$

The share of the dividends to pay is assumed to be fixed for all the banks at the level of 0.5 of the net profit.

After all the markets are cleared, the banks update their expectations. Each bank forms the expectations for the central bank interest rate $i^e_{cb,t-1}$, the credit $i^e_{Cr,t-1}$, the deposit market $i^e_{Dp,t-1}$, the interbank market $i^e_{Bb,t-1}$; the prices for the labor market $p^e_{L,t-1}$, the capital market $p^e_{K,t-1}$, the governments bonds market $p^e_{Bg,t-1}$. The rule for updating the expectations follows the scheme developed in the previous chapter, with details given in Appendix A.2.

The decision choice d from the domain D^b is made using the ADP algorithm with simplification used in previous Chapter 2 and described in Appendix A.3.

4.3.2 Humans

A household in the current model consists of one person who lives infinitely. From now on, the term human will be used to denote a household.

4.3.2.1 Human constraints and goals

The utility function assumes the usual form of the expected utility:

$$U\left(\{cons_i, k_i, l_i\}_{i=t}^{\infty}\right) = \mathbb{E}_t\left(\sum_{i=t}^{\infty} \beta^i u\left(cons_i, k_i, l_i\right)\right) \tag{4.24}$$

with the period utility given by:

$$\begin{aligned} U(cons_t, k_t, l_t) &= \theta_0 log\left(1 + cons_t\right) \\ &+\theta_1 log\left(1 + k_t\right) \\ &+\theta_2\left(1 - l_t\right) \end{aligned} \tag{4.25}$$

where c_t is the consumption good, k_t are the services provided from the capital good, and l_t are the labor services supplied by the agent.

In each period, humans are subject to a sequence of budget constraints in the spirit of the previous work. These constraints are given below in equations (4.26), (4.27), (4.28), (4.30), (4.30), (4.31).

The capital services are defined by the stock of capital goods owned by the human:

$$k_t \;=\; K_{t:0} \tag{4.26}$$

The purchases of goods $q_{C,t:C}$ and capital $q_{K,t:K}$ have to be financed by the available money balances, no credit purchases are allowed:

$$q_{C,t:C} \cdot p_{C,t:C} \leq M_{t:C-1} \tag{4.27}$$

and

$$q_{K,t:K} \cdot p_{K,t:K} \leq M_{t:K-1} \tag{4.28}$$

The money balances follow the update rule given below:

$$\begin{aligned} M_{t+1:0} \;=\; M_{t:0} \;+\; w_{t:Clear} \\ -\; i_{Cr,t:Clear} \;-\; c_{Cr,t:Clear} \\ +\; i_{Dp,t:Clear} \;-\; d_{Dp,t:Clear} \\ +\; tr_{t:Tr} \;-\; tax_{t:Tax} \;+\; \sum_{v \in V} div_{t:Div} \\ -\; q_{C,t:C} \cdot p_{C,t:C} \;-\; q_{K,t:K} \cdot p_{K,t:K} \\ +\; c_{t:Cr} \;-\; d_{t:Dp} \end{aligned} \tag{4.29}$$

During the clearing stage at the beginning of each period, the human receives the promised wages $w_{t:Clear}$; the promised interest and body payments on the deposit contracts $i_{Dp,t:Clear}$, $d_{Dp,t:Clear}$; the transfers from the government $tr_{t:Tr}$; the dividends from the other agents $\sum_{v \in A} div_{t:Div}$, where A is the set of all the agents in the economy. He also pays the promised interest and body payments on the credits $i_{Cr,t:Clear}$, $c_{Cr,t:Clear}$ and the taxes $tax_{t:Tax}$. When the consumption and capital goods market function, the agent purchases $q_{C,t:C}$ of consumption goods and $q_{K,t:K}$ of capital goods and makes the corresponding payments $q_{C,t:C} \cdot p_{C,t:C}$ and $q_{K,t:K} \cdot p_{K,t:K}$. When the credit and deposit markets function, the agent borrows the amount of money $c_{t:Cr}$ and makes the deposits in the amount of $d_{t:Dp}$.

Given the stock of capital $K_{t:0}$ and the new purchases of capital goods $q_{K,t:K}$, the capital accumulation equation is:

$$K_{t+1:0} = (1 \;-\; \delta_{dep})\,(K_{t:0} + q_{K,t:K}) \tag{4.30}$$

The human consumes all the consumption goods that are available at the end of the period, so

$$cons_t \;=\; q_{C,t:C} \tag{4.31}$$

All the relevant variables, such as the consumption, the amount of goods, etc., are subject to the non-negativity constraints.

4.3.2.2 Decision domain and transformation functions for humans

For each human, the decision domain D^h is defined as follows.

$$D^h \;=\; \Theta^L \otimes \Omega^L \otimes \Theta^K \otimes \Theta^C \otimes \Theta^{Cr} \otimes \Theta^{Dp} \tag{4.32}$$

where:

- the elements of $\Theta^L = \{\theta^C_1, \ldots, \theta^C_C\}$ satisfy $0 \le \theta^C_1 < \ldots < \theta^C_C \le 1$
- the elements of $\Omega^L = \{\omega^L_1, \ldots, \omega^L_L\}$ satisfy $0 < \omega^L_1 < \ldots < \omega^L_L$
- the elements of $\Theta^K = \{\theta^K_1, \ldots, \theta^K_K\}$ satisfy $0 \le \theta^K_1 < \ldots < \theta^K_K \le 1$
- the elements of $\Theta^C = \{\theta^C_1, \ldots, \theta^C_C\}$ satisfy $0 \le \theta^C_1 < \ldots < \theta^C_C \le 1$
- the elements of $\Theta^{Cr} = \{\theta^{Cr}_1, \ldots, \theta^{Cr}_{Cr}\}$ satisfy $0 \le \theta^{Cr}_1 < \ldots < \theta^{Cr}_{Cr}$
- the elements of $\Theta^{Dp} = \{\theta^{Dp}_1, \ldots, \theta^{Dp}_{Dp}\}$ satisfy $0 \le \theta^{Dp}_1 < \ldots < \theta^{Dp}_{Dp} \le 1$

θ^L is the share of the human capital endowment to offer on the labor market. The wage offer for the labor market equals $\omega^L w^e_{t-1}$, where w^e_{t-1} is the expected clearing labor wage. θ^K is the share of the money balances $M_{t:K-1}$ to spend on the capital. θ^C is the share of the money balances $M_{t:C-1}$ to spend on the consumption goods. θ^{Cr} is the share of the assets $Assets_{t:Cr-1}$ to borrow, with $i^{bid}_{Cr,t:Cr} = i^e_{Cr,t-1} + 2\sigma^{2,e,Cr}_{t-1}$ being the bid price. θ^{Dp} is the share of the money balances $M_{t:Dp-1}$ to deposit, with $i^{ask}_{Dp,t:Dp} = i^e_{Dp,t-1}$ being the ask price for the deposit market.

The bids and asks for the markets are formed in the same way as described in Chapter 2 in Sections 2.3.2, 2.3.3.

After the capital market has been cleared, the new purchases $q_{K,t:K}$ will be incorporated into the capital stock in the next period according to equation (4.30).

The human forms the expectations for the credit $i^e_{Cr,t-1}$, the deposit market $i^e_{Dp,t-1}$, the prices for labor market $p^e_{L,t-1}$, the capital market $p^e_{K,t-1}$, the goods market $p^e_{C,t-1}$. The rule for updating the expectations follows the scheme previously developed in Chapter 2, with details given in Appendix A.2.

4.3.3 Firms

There are two types of firms. Those of the first type produce consumption goods, and those of the second type produce capital goods. A typical setup for this part of economy is to have producers of intermediate goods that are subject to monopolistic competition, and after that the final goods producers that combine intermediate goods to produce consumption (or final) goods for sale to consumers. In this setup, the choice was made to have separate capital and consumption goods producers. Capital is essential in this economy to facilitate intermediation of the financial resources. The consumption goods are goods that are not storable by humans and that enter their utility functions.

4.3.3.1 Firm constraints and goals

Firms try to maximize their expected profits

$$\max_{d \in D^b} E_{v,t} \sum_{r=t}^{\infty} \mu^{r-t} \left[\Pi_t \left(\mathbf{X}_r, d, \mathbf{W}_r \right) \right] \tag{4.33}$$

subject to no-bankruptcy conditions

$$M_{t:Bankruptcy} \geq 0 \tag{4.34}$$

The resource accumulation constraints are given by equations (4.36),(4.36),(4.38).

The money balances M_t follow the update rule given below:

$$\begin{aligned} M_{t+1:0} \ = \ M_{t:0} \ - \ w_{t:Clear} \\ - \ i_{Cr,t:Clear} \ - \ c_{Cr,t:Clear} \\ + \ i_{Dp,t:Clear} \ - \ d_{Dp,t:Clear} \\ - \ tax_{t:Tax} \ - \ div_{t:Div} \\ + \ q_{G,t:G} \cdot p_{G,t:G} \\ - \ q_{K,t:K} \cdot p_{K,t:K} \\ + \ c_{t:Cr} \ - \ d_{t:Dp} \end{aligned} \tag{4.35}$$

During the clearing stage at the beginning of each period, the firm pays the promised wages $w_{t:Clear}$; the interest and body payments on credits $i_{Cr,t:Clear}$, $c_{Cr,t:Clear}$; the dividends $div_{t:Div}$; the taxes $tax_{t:Tax}$. On the other hand, the firm receives the promised interest and body payments on deposit contracts $i_{Dp,t:Clear}$, $d_{Dp,t:Clear}$. When consumption and capital goods markets function, the firm sells $q_{C,t:C}$ of consumption goods if it produces consumption goods, or $q_{K,t:K}$ of capital goods, if it produces capital goods. The firm receives the corresponding payments $q_{C,t:C} \cdot p_{C,t:C}$ or $q_{K,t:K} \cdot p_{K,t:K}$, payment is $q_{G,t:G} \cdot p_{G,t:G}$ with $G \in \{C, K\}$. Besides that, the firm buys the investment in the amount of $q_{K,t:K}$ and pays $q_{K,t:K} \cdot p_{K,t:K}$. When the credit and deposit markets function, the agent borrows money $c_{t:Cr}$ and makes deposits $d_{t:Dp}$.

Given the stock of capital $K_{t:0}$ and the new purchases of capital goods $q_{K,t:K}$, the capital accumulation equation is:

$$K_{t+1:0} = (1 \; - \; \delta_{dep}) \, (K_{t:0} + q_{K,t:K}) \tag{4.36}$$

The production function for the firm producing goods of the type G is:

$$q_{G,t:Production} \;\; = \;\; F^{G} \, (l_t, k_t) \;\; = \;\; A^{G} l_t^{\alpha^G} k_t^{\beta^G} \tag{4.37}$$

parameters A^G, α^G, β^G are given in the Appendix C in the Table C.1.

Given production $q_{G,t:Production}$, the stock of goods available for sale is updated in the following way:

$$q^{stock}_{G,t:Production+1} \;\; = \;\; q^{stock}_{G,t:Production-1} \; + \; q_{G,t:Production} \tag{4.38}$$

The period profit is defined as follows:

$$\begin{aligned} \Pi_t \;\; &= \;\; i_{c,t} \left(c_{t:Cr}, i_{c,t}, \mathbf{c}_{t-1} \right) \\ &- \; i_{d,t} \left(d_{t:Dp}, i_{d,t}, \mathbf{dp}_{t-1} \right) \\ &+ \; pq_{d,t} \left(p_{G,t:G}, q_{G,t:G} \right) \\ - \; cost_t \left[w_t \left(l_t, \mathbf{w}_{t:L}, \mathbf{hk}_{t-1} \right), dep_t \left(k_t, \mathbf{p}_{K,t:K}, \mathbf{k}_t \right), \mathbf{cost}_{t-1} \right] & \end{aligned} \tag{4.39}$$

In equation (4.39), $i_{c,t}$ is the current period expenses on the credit. They include the interest payments that are acquired in the current period from the extended loans c_t, and are generally equal to $c_t \cdot i_{c,t}$, and include the interest payments from the loans extended in the previous periods that are still outstanding. Here $\mathbf{c}_{t-1}$ includes the information about all the loans extended in periods r, where $0 \leq r < t$.

$i_{d,t}$ is the current period income from the deposits. This includes the interest income that is acquired in the current period from made deposits d_t, and which is generally equal to $d_t \cdot i_{d,t}$, and includes the interest payments on the deposits accepted in the previous periods that are still outstanding. Here $\mathbf{d}_{t-1}$ includes the information about all the deposits accepted in periods r, where $0 \leq r < t$.

$pq_{d,t}$ is the current period income from the sale of goods. It is defined by the current sale of produced goods $q_{G,t:G}$ and is equal to $p_{G,t:G} \cdot q_{G,t:G}$.

$cost_t$ are the costs of production. They are calculated as an average cost of sold goods. The average cost is calculated in the following way. At every period, the current period expenses on labor w_t and the current depreciation of capital dep_t (defined below) are added to the total cost of goods currently held in stock. The new average cost is equal to the total costs divided by the total amount of goods, both held in inventory and produced in the current period. Here $\mathbf{cost}_{t-1}$ is the previous period average cost and inventories.

w_t are the current period expenses on labor. They are defined by the currently employed amount of labor l_t and the promised wages $\mathbf{w}_{t:L}$. Here $\mathbf{hk}_{t-1}$ includes the information about all the labor contracts signed in previous periods r, with $0 \leq r < t$, that are still active.

dep_t is the current period expenses on capital. They are defined by the current depreciation of capital, which is a function of the current capital stock k_t and the cost of acquiring capital $\mathbf{p}_{k,t:K}$. Here $\mathbf{k}_{t-1}$ includes the information about all the capital purchases in previous periods r, with $0 \leq r < t$.

All the relevant variables, such as the consumption, the amount of goods, etc., are subject to non-negativity constraints.

4.3.3.2 Decision domain and transformation functions for firms

The firm's choice domain is defined as follows:

$$D^f = \Omega^L \otimes \Theta^L \otimes \Theta^K \otimes \Omega^G \otimes \Theta^{Cr} \otimes \Theta^{Dp} \tag{4.40}$$

where:

- the elements of $\Theta^L = \{\theta^L_1, \ldots, \theta^L_L\}$ satisfy $0 \leq \theta^L_1 < \ldots < \theta^L_L \leq 1$

- the elements of $\Omega^L = \{\omega_1^L, \ldots, \omega_L^L\}$ satisfy $0 < \omega_1^L < \ldots < \omega_L^L$
- the elements of $\Theta^K = \{\theta_1^K, \ldots, \theta_K^K\}$ satisfy $0 \leq \theta_1^K < \ldots < \theta_K^K \leq 1$
- the elements of $\Omega^G = \{\omega_1^G, \ldots, \omega_G^G\}$ satisfy $0 < \omega_1^G < \ldots < \omega_G^G$
- the elements of $\Theta^{Cr} = \{\theta_1^{Cr}, \ldots, \theta_{Cr}^{Cr}\}$ satisfy $0 \leq \theta_1^{Cr} < \ldots < \theta_{Cr}^{Cr}$
- the elements of $\Theta^{Dp} = \{\theta_1^{Dp}, \ldots, \theta_{Dp}^{Dp}\}$ satisfy $0 \leq \theta_1^{Dp} < \ldots < \theta_{Dp}^{Dp} \leq 1$

where θ^L is the share of the capital stock $K_{t:L-1}$ to be hired as labor. The wage to offer equals to $\omega^L w_{t-1}^e$, where w_{t-1}^e is the expected clearing labor wage. θ^K is the target share of the assets to have as capital.

The desired quantity to buy from the capital market equals

$$q_{t:K}^K \;=\; \theta^K Assets_{t:K-1}/p_{t-1}^e - K_{t:K-1} \tag{4.41}$$

The actual bid is formed in the following way: for each price p_K, the quantity to buy equals to $q_{t:K}^{bid} = [\min(M_{t:K-1}/p_K, q_{K,t:K})]$.

The share of the stock of goods to sell is $\theta^G \;=\; 1$ and the price to sell that is formed is $p_{G,t:G}^{ask} \;=\; \omega^G p_{G,t:G}^e$, where G stands for both the capital and consumption goods markets, depending on the firm's specialization.

For the borrowing decisions, θ^{Cr} is the share of the assets to borrow. The bid for the credit market would have $\theta^{Cr} Assets_{t:Cr-1}$ as a target value, and the bid price would be $i_{Cr,t:Cr}^{bid} \;=\; i_{Cr,t-1}^e \;+\; 2\sigma_{t-1}^{2,e,Cr}$. For the lending decisions, $\theta^{Dp} Assets_{t:Dp-1}$ is the ask quantity, with $i_{Dp,t:Dp}^{ask} \;=\; i_{Dp,t-1}^e$ being the ask price for the deposit market.

The share of the dividends to pay θ^{div} is assumed to be fixed for all the agents at the level of 0.25 of the net profit. The assets for firm v at time t are the sum of the money balances M_t and the valuations for the consumption and capital goods $(\sum_{i\in\{K,C\}} q_{i,t} p_{i,t-1:i})$.

$$Assets_{v,t:r} \;=\; M_{v,t:r} \;+\; \sum_{i\in\{K,C\}} q_{i,t:r} p_{i,t-1:i} \tag{4.42}$$

The bids and asks for the markets are formed in the same way as described in Chapter 2 in Sections 2.3.2, 2.3.3.

The firm forms expectations for the credit $i^e_{Cr,t-1}$, the deposit market $i^e_{Dp,t-1}$, the prices for the labor market $p^e_{L,t-1}$, the capital market $p^e_{K,t-1}$, the goods market $p^e_{C,t-1}$. The rule for updating the expectations follows the scheme previously developed in Chapter 2, with details given in Appendix A.2.

4.3.4 Other Agents that Follow Fixed Rules

The government in this model takes actions according to fixed rules.

Every few periods $f\_bg \; = \; 6$, the government issues bonds at the fixed quantity $\theta^{Bg} \; = \; 100$ and the fixed price $\omega^{Bg} \cdot fv_{bg} \; = \; 0.9 \cdot 10.0$, with $fv_{bg} \; = \; 10.0$ being the face value of the bonds. The announced interest rate is also fixed at the value of $i_{bg} \; = \; 0.04$. The length of a contract is fixed at $c\_length\_bg \; = \; 6$.

Every period it also decides on the amount of transfers to the humans. The transfers are proportional, with the proportionality rate equal to $\theta^{Tr} \; = \; 0.5$, to the money balances $M_{g,t}$ held at the central bank, adjusted for the expected payments on the outstanding bonds $\mathbf{bg}_t$.

The money balances are increased when a profit from the central bank is transferred or when taxes are paid. They are decreased when obligations on bonds are paid and transfers are made.

$$M_{g,t} \; = \; M_{g,t-1} \; + \; \sum_{v \in A} div_{v,t:Div} \; + \; \sum_{v \in A} tax_{v,t:Tax} \; - \; \sum_{v \in A} tr_{v,t:Tr} \; - \; i_{bg,t} \left(i_{bg}, \mathbf{bg}_t \right) \quad (4.43)$$

The central bank also follows a set of fixed rules. The central bank sets the fixed interest rate $\omega_{Cb} \; = \; 0.1$ on the standing facilities. The amount of credit that each bank v could receive from the central bank could not be more than 5% of the total $Capital_{v,t}$.

The central bank also implements purchases on the market for the government bonds in the fixed amounts $q^{Cb}_{Bg:primary}$, $q^{Cb}_{Bg:secondary}$. The amount of purchases is one of the treatment factors in the model. If no QE policy is implemented, then the central bank do not participate on the market for the government bonds and $q^{Cb}_{Bg:primary} = 0.0$, $q^{Cb}_{Bg:secondary} = 0.0$. If QE1 is implemented, then $q^{Cb}_{Bg:primary} = 20.0$ for the primary market, and $q^{Cb}_{Bg:secondary} = 0.0$ for the secondary market. In the case of QE1, the central bank participates only on the primary government bonds market. If QE2 is implemented, then $q^{Cb}_{Bg:primary} = 20.0$ for the primary market, and $q^{Cb}_{Bg:secondary} = 20.0$ for the secondary market. In the case of QE2, the central bank participates on both the primary and secondary markets. The bid price for the market of the government bonds p^{bid}_{Bg} is set to infinity.

The central bank receives the period profit Π_t from operations on the government bonds markets $i_{bg,t}\left(bg_{t:Bg}, i_{bg,t}\right)$, and from the use of standing facilities $i_{cb,t}\left(cb_{t:Cb}, i_{cb,t}\right)$, both of them in the form of interest payments on the contracts.

$$\Pi_t \;=\; i_{cb,t}\left(cb_{t:Cb}, i_{cb,t}\right) \;+\; i_{bg,t}\left(bg_{t:Bg}, i_{bg,t}\right) \tag{4.44}$$

All the profits from operations are transferred back to the government.

4.3.5 Institutional Structure

All the markets in the simulations are modeled as competitive markets, with each agent submitting a bid and an ask, and the market solving for the equilibrium price and quantity. After that, the bids and asks are fulfilled with the equilibrium price. The general description of these markets is given in Section 2.2.3. Specific details for each market are shown below.

4.3.5.1 Market for human capital

The market for contracts for human capital has humans, firms and banks as active agents. Each of them makes decision regarding their bids and asks at the time of decision making and submits them to be cleared by the market. Market open for trade after current contracts has expired which happens every c_length_Hk periods.

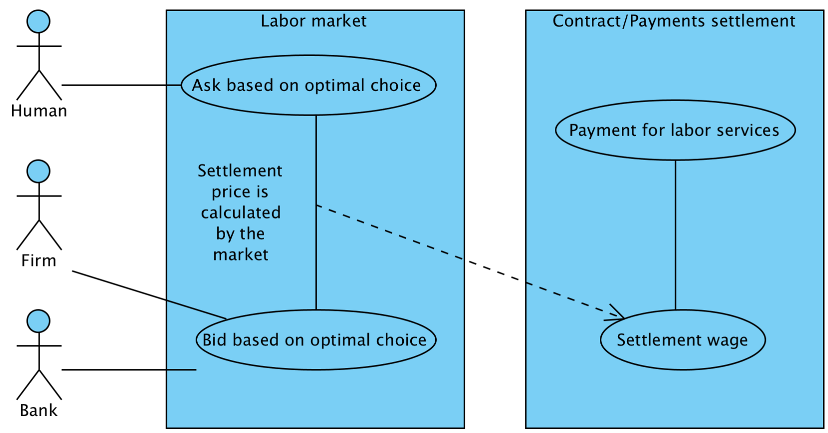

Figure 4.3 Description of the labor market functioning

Each market bid(ask) starts as a decision of an agent. Typically such decisions are made in terms of the number of the human capital units that the agent is willing to supply or buy from the market. Each agent presents bids and asks in the form of a simple spline demand/supply curve with one node. For bids, this point represents the maximum price agent is willing to pay, and for asks the minimum price that the agent is willing to accept.

A contract includes the length c_length_Hk, wage payments per period, the issuer and the holder. Contracts are automatically terminated when they end.

4.3.5.2 Market for goods

For the goods market, the main agents are humans and firms. Humans need consumption goods as a part of their utility. They also need some leisure and services provided by capital goods. Besides that, firms and banks need capital that is provided by capital producing firms.

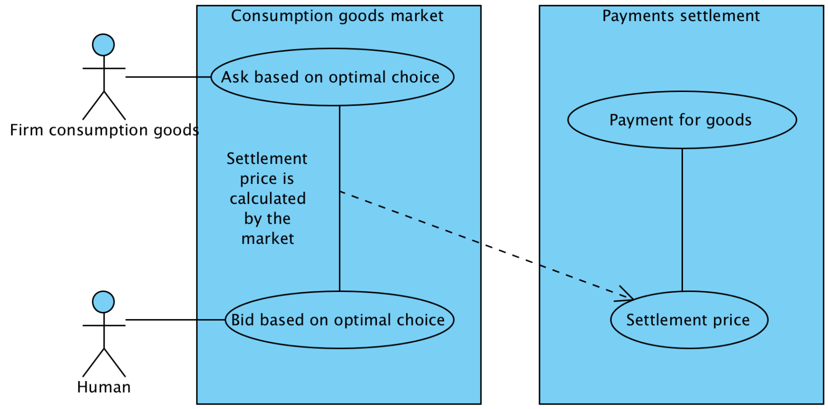

Figure 4.4 Description of the consumption goods market functioning

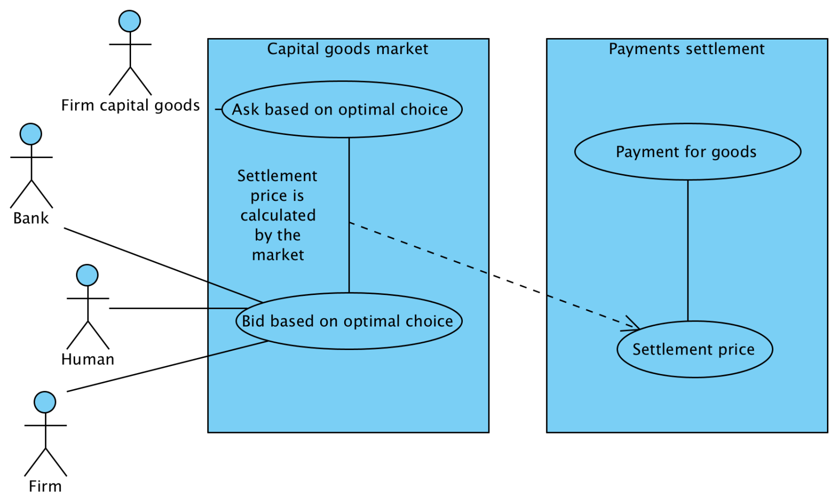

Figure 4.5 Description of the capital goods market functioning

Bids are submitted in the form of a potential demand curve, the range of prices is restricted to the interval $\left[\theta_{min,pG} * p^e_{G,t-1:G}, \theta_{max,pG}\right]$. Where $p^e_{G,t-1:}$ is expected by an agent price for the market G. This was done to aid stabilization of numerical calculations. Asks are submitted in the form of supply curves with one switching point. Ones market clears and clearing price is calculated, the delivery of goods is instantaneous.

4.3.5.3 Interbank market

The interbank market is a short-term market for funds at accounts at the central bank for this economy. In this market, each bank submits the amount to sell/buy and cut-off interest rates.

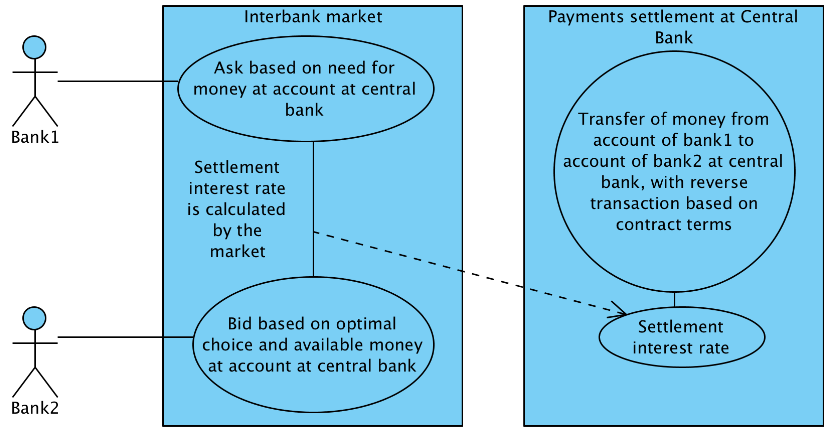

Figure 4.6 Description of the interbank market functioning

The bids and asks are described in 4.3.1. After clearing, the interest rate is decided by the market, the appropriate contracts are signed, and after that the transfer of money takes place.

The duration of a contract is $c\_length\_Bb = 1$, so contracts expire when the next period starts. Contracts include the length c_length_Bb, the interest per period and the body payments, the issuer and the holder. Contracts are automatically terminated when they end.

4.3.5.4 Credit market

On the credit market, the main participants are humans, firms as buyers and banks as sellers of credit contracts. Bids and asks are described in 4.3.1, 4.3.3, 4.3.2.

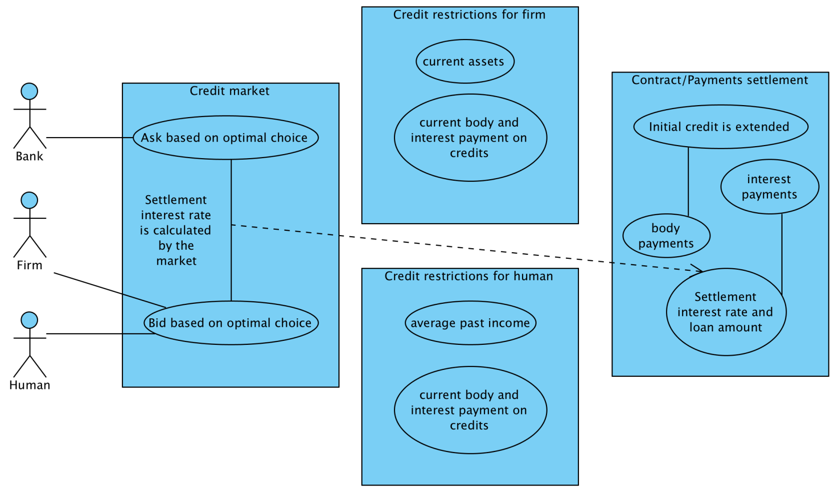

Figure 4.7 Description of the credit market functioning

On the top of calculating the settlement price, there is also a requirement on the amount of the individual credits that could be extended. Every bank has a limit on how much money could be loaned to each individual agent, as well as a restriction on the total outstanding credit. Individual restrictions are defined by the potential of an agent (earning for humans and assets for firms). This potential is corrected for current payments on outstanding loans in the following way.

For humans:

$$cout^h_{max} = \theta_{c^h} * \min\left(cout^{bid}, \overline{income} - cout^{h,body}_t - cout^{h,i}_t\right) \tag{4.45}$$

where $\overline{income}$ is the average past income of the human, which includes the interest, the labor and financial income in the form of dividends. $cout^{h,body}_t$ are the current credit payments on outstanding loans, $cout^{h,i}_t$ are the current interest payments on outstanding loans. θ_{c^h} is the parameter of the bank decision process and is defined in Appendix C in Table C.8.

For firms:

$$cout^f_{max} = \theta_{cout^f} * \min\left(c^{bid}, Assets_t - cout^{f,body}_t - cout^{f,i}_t\right) \tag{4.46}$$

where $Assets_t$ are assets of the firm. $cout_t^{f,body}$ are the current credit payments on outstanding loans, $cout_t^{f,i}$ are the current interest payments on outstanding loans. θ_{cf} is a parameter of bank decision process that is defined in Appendix C in Table C.8.

In the current version, there is no check for the history of default.

Once the market clears, the contracts are signed. Contracts include the length *c_length_Cr*, the interest per period and the body payments, the issuer and the holder. Contracts are automatically terminated when they end.

4.3.5.5 Deposit market

On the deposit market, the main participants are humans, firms as sellers and banks as buyers of deposit contracts. Details on bids and asks are described in Sections 4.3.1, 4.3.3, 4.3.2.

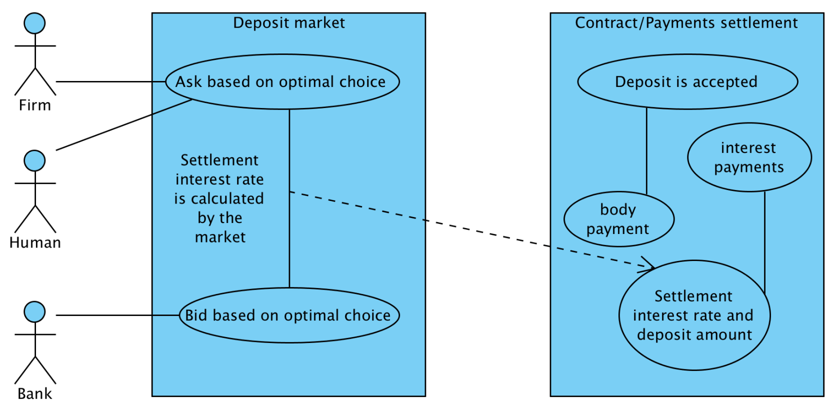

Figure 4.8 Description of the deposits market functioning

Humans and firms decide upon the share of assets to have invested as deposits, while banks have imposed restrictions on the amount of deposit they could accept.

Once the market clears, the contracts are signed. Contracts include the length *c_length_Dp*, the interest per period and body payments, the issuer and the holder. Contracts are automatically terminated when they end. The length of a deposit contract is *c_length_Dp* and is given in Appendix C in Table C.9.

4.3.5.6 Government bonds market

The government issues bonds on the primary market. A new issue is offered after the previous bond issue has expired. An ask comes from the government and has q, p_{min}, i, where q^{ask}_{Bg}, p^{ask}_{Bg}, i^{ask}_{Bg} are taken from the preset parameters and are described in Section 4.3.4. Bids come from the central bank, which prioritizes the initial purchases over the secondary market and sets p^{bid}_{Bg}, q^{bid}_{Bg} equal to the preset parameters which are described in Section 4.3.4, and banks which do not distinguish between the primary and secondary markets and submit bid prices and quantities as defined in Section 4.3.1. For the secondary market, bids and asks are submitted by banks based on the desired share of the government bonds in assets as described in Section 4.3.1 and by the central bank based on the desired quantity of the government bonds in the portfolio as described in Section 4.3.4.

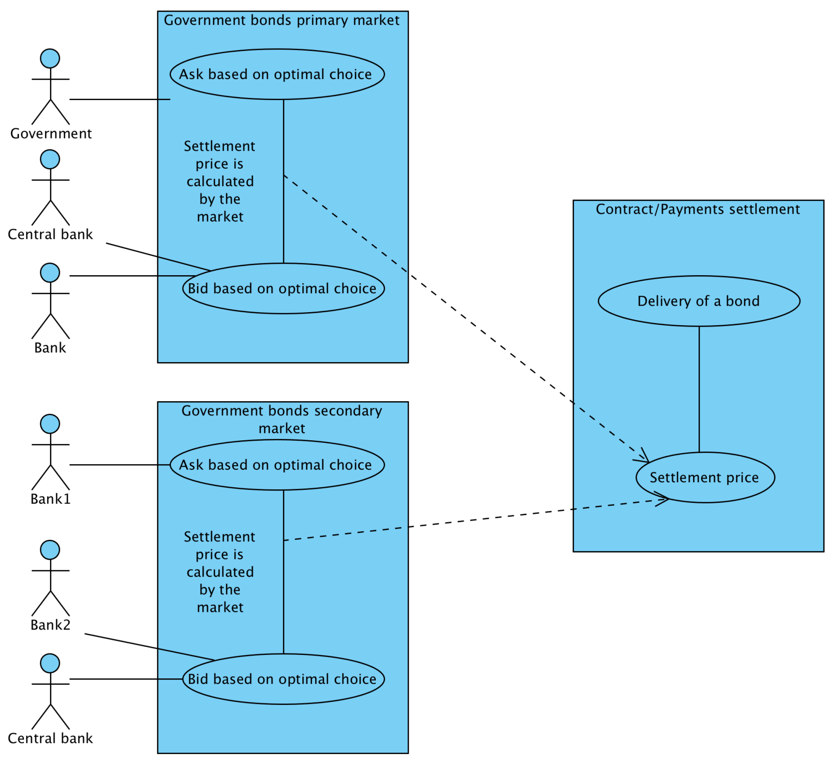

Figure 4.9 Description of the government bonds market functioning

Once the market clears, the bonds are transferred. Bond contracts include the length *c_length_Bg*, the interest per period and body payments, the issuer and the holder. Contracts are automatically terminated when they end. The length of a contract *c_length_Bg* is given in Appendix C in Table C.9.

4.3.5.7 Central bank standing facilities

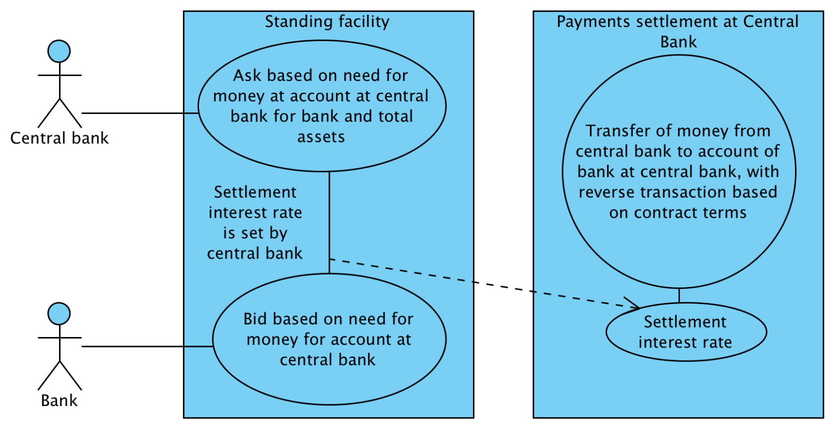

Figure 4.10 Description of the central bank standing facility

Details on the functioning of this market are described in Section 4.3.4. Standing facilities are available to banks and provide them with back up money resources in the case when their money balances become negative during their daily interactions.

4.3.5.8 Accounting

In general, the accounting is modeled to be as close as possible to the International Accounting Standards (IAS).

For firms, the average method for cost calculations is used. Adjustments for zero sales or production in the period are included. Depreciation is accounted as a cost for the period when it happened. Another part of the cost are wage payments for active contracts, which are accounted in the period when the labor services are provided.

Income is generated by selling produced goods and from investing spare money balances in the deposit instruments (realized). The profit is the difference between the income and the costs in agreement with the usual definition.

For banks, the income is comprised of the interest on the extended credit, the interest from bond holding, the interest from interbank loans. The interest income is accounted in the realized form to allow for bankruptcy accounting. Expenses are formed by the interests paid on deposits, on interbank loans, on standing facilities, the depreciation and the labor expenses. The profit is the difference between the income and the expenses. The current income/loss from holding government bonds is added to the profit.

For the central bank, the income is the interest paid on standing facilities and on the holdings of the government bonds. The profit is distributed back to the government.

For the government, accounting consists of tracking transfers to humans and taxes paid by all agents.

4.3.5.9 Taxes and dividends

In the current model, banks and firms pay the profit tax that equals a certain share of the profit, and humans pay the income tax, which is a certain share of the labor income they receive.

These shares are fixed at the levels of $tax\_f = 0.1$ for firms, $tax\_b = 0.1$ for banks, $tax\_h = 0.05$ for humans and $tax\_cb = 1.0$ for the central bank.

The government uses tax payments to pay out bonds and distribute transfers to humans.

Dividends are paid as a certain share of the net profit (excluding tax payments) acquired by banks and firms. Taxes are paid with the periodicity f_taxes, as described in section 4.2.3. Dividends are paid with periodicity f_div as described in section 4.2.3.

4.3.5.10 Payment system

The payment system is one of the invisible, but important, parts of the model. All payment orders go through banks that have accounts of the agents (with the exception of the government, which has an account with the central bank). If a payment order involves payments for accounts at the same bank, they are cleared in the bank; otherwise, the payment order goes to the central bank for clearing. The corresponding amounts are subtracted/added to bank accounts involved in the transaction. Each participating agent gets information about the payment. If the payment involves government, then it is processed as if the counterparty was a bank with an account at the central bank.

A general depiction of the payment system is given below.

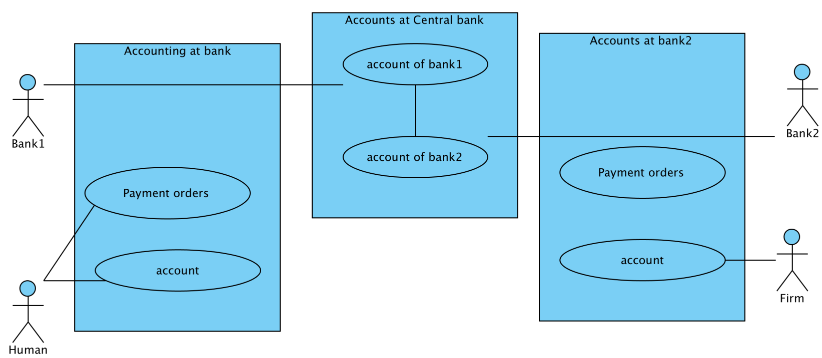

Figure 4.11 Description of the payment system

4.3.5.11 Bankruptcy

When a firm or a bank has no money to pay the current payment order, a bankruptcy is initiated. An agent is marked as a bankrupt, and all future payments on his contracts are stopped. The agent is restored to the active state after a fixed number of periods. Therefore, a bankruptcy is handled as shown below.

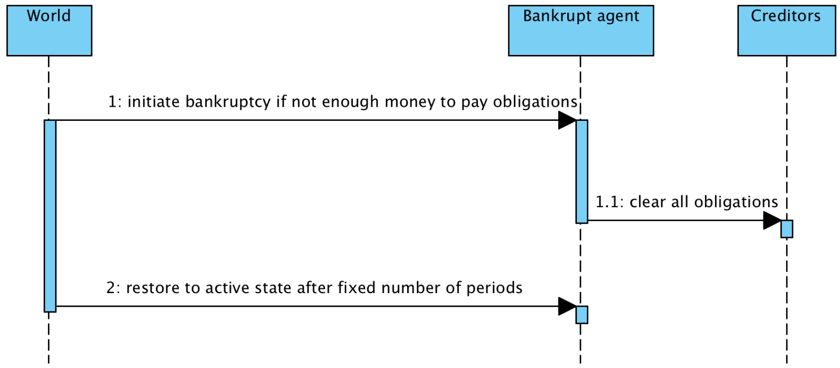

Figure 4.12 Description of the bankruptcy procedure

If the bankruptcy condition, defined for firms by equation (4.34) and for banks by equation (4.2), is met, then the bankruptcy procedure is implemented.

It is worth noting that in this model agents as assumed to be bankrupt if they satisfy the requirements for the cash flow insolvency without any court of law legal orders. Balance sheet insolvency does not trigger bankruptcy procedures in this version.

Bankruptcies are implemented in two different ways. These implementations represent one of the treatment factors in the model. The first way to implement bankruptcy $BankruptcyNoRestock$ is to remove agents for 4 periods from the economy. During this time, no contract payments are made and due contracts are allowed to expire. The second way called $BankruptcyRestock$ removes agents from the economy for the same amount of periods, but after that restores their money balances and capital for banks, capital goods stock for firms (if it became negative) to the level of the priors defined in Appendix C in Table C.10.

4.4 Algorithms Used in the Model

The expectation formation rules are similar to those used for the EO-ADP agents in Chapter 2. Each agent tracks a subset of all the market prices and updates the expected mean and variance for them. For each price, each agent has prior expectations, receives the same information from the market as everybody else about the realized clearing price and recalculates the average and variance accordingly.

Each price $w_{t,i}$ from the set of all prices $\boldsymbol{w_t} = [w_{t,1}, w_{t,2}, \cdots, w_{t,n-1}, w_{t,n}]'$ is assumed to be formed independently and to follow a normal distribution. So $\boldsymbol{w_t} \sim \boldsymbol{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma_t})$, with

$$\boldsymbol{\Sigma_t} = \begin{bmatrix} \sigma^2_{t,1} & 0 & \cdots & 0 & 0 \\ 0 & \sigma^2_{t,2} & 0 & \ddots & 0 \\ \vdots & 0 & \cdots & 0 & \vdots \\ 0 & \ddots & 0 & \sigma^2_{t,n-1} & 0 \\ 0 & 0 & \cdots & 0 & \sigma^2_{t,n} \end{bmatrix} \tag{4.47}$$

Thus, $\boldsymbol{w_t} = [z_{t,1}, z_{t,2}, \cdots, z_{t,n-1}, z_{t,n}]'$ with each $z_{t,i}$ being independent, and $z_{t,i} \sim N(\mu_i, \sigma_i^2)$. These expected distributions are updated according to the process described in Appendix A.2. Also, as mentioned before, each agent v uses the EO-ADP solution method to chose the optimal decision d from the decision domain D^v. This algorithm is described in Appendix A.3.

4.5 Design of the Sensitivity Analysis and Dynamics of the Model

4.5.1 Design of the Sensitivity Analysis

The initial parameter values for the expectations and the stock of goods, as well as the parameters of the production function, are described in Appendix C. The model was run in three modes corresponding to the types of the policies implemented by the central bank. The difference between the tested policies is the amount of the intervention by the central bank on the primary and the secondary markets for the government bonds. The details of this treatment are described in Section 4.3.4 and are also summarized below.

In the case of no QE ($noQE$) policies, the relevant central bank decision parameters were set as follows:

Table 4.1 Parameter values for the case of noQE policy

Parameter	**Value**
$q^{Cb}_{Bg:primary}$	0.0
$q^{Cb}_{Bg:secondary}$	0.0

In the case of the $QE1$ policies, the relevant central bank decision parameters were set as follows:

Table 4.2 Parameter values for the case of QE1 policy

Parameter	**Value**
$q^{Cb}_{Bg:primary}$	20.0
$q^{Cb}_{Bg:secondary}$	0.0

In the case of the $QE2$ policies, the relevant central bank decision parameters were set as follows:

Table 4.3 Parameter values for the case of QE2 policy

Parameter	**Value**
$q^{Cb}_{Bg:primary}$	20.0
$q^{Cb}_{Bg:secondary}$	20.0

These policies $noQE$, $QE1$, $QE2$ were tested for two possible institutional frameworks $BankruptcyNoRestock$ and $BankruptcyRestock$ described in Section 4.3.5.11. The results of simulations are presented below.

4.5.2 Dynamics of the Model

4.5.2.1 Results for the case of the simplified bankruptcy arrangements

In the case of the simplified treatment of bankruptcy *BankruptcyNoRestock*, the following results for the main macroeconomic indicators were produced.

Each indicator x given below, unless specified otherwise, is calculated by averaging over the simulation runs $NRuns$ in the following way:

$$\bar{x}_t = \frac{\sum_{r=1}^{NRuns} x_{t,r}}{NRuns} \tag{4.48}$$

The level of the credits, which is equal to the sum of all outstanding loans at time t in the economy, is presented below.

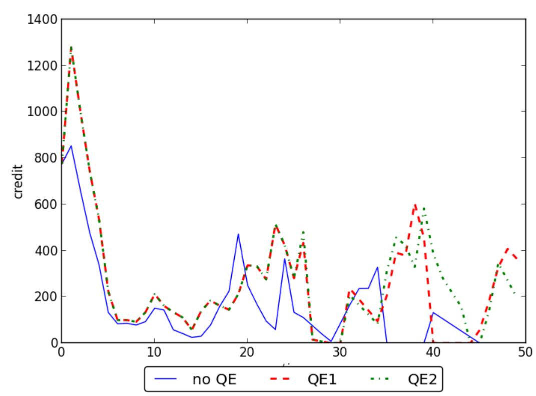

Figure 4.13 Simulation results for the credit levels, the case of simplified bankruptcy

In this case, the agents in the economy take on excessive risks at the beginning, and subsequently go bankrupt. Bad expectations are formed and solidified, leading to an eventual decrease in the employment, production and credit activity. This situation can be described as a wrong risk assessment on the part of agents, since the economy is too complicated for them to form correct expectations.

The employment in period t is defined as a total number of active labor contracts. The employment level is deteriorating during the simulations.

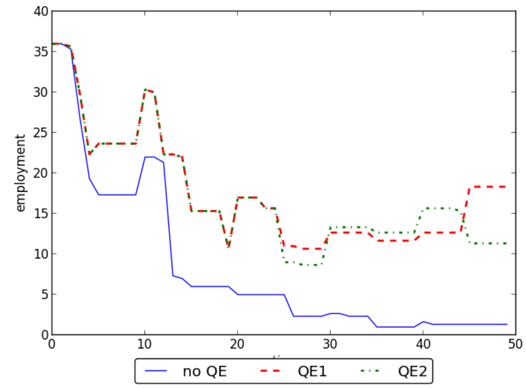

Figure 4.14 Simulation results for the employment levels, the case of simplified bankruptcy

The level of deposits is defined by the total amount of outstanding deposits in the economy in period t. The dynamics of this parameter are presented below:

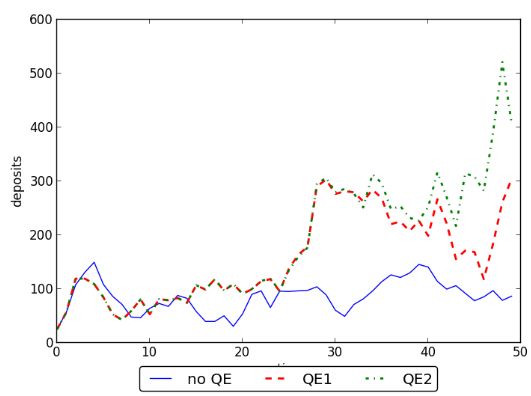

Figure 4.15 Simulation results for the deposit levels, the case of simplified bankruptcy

While QE makes a difference in terms of the government bonds levels and deposits, it does not influence the long-term economic results, as can be seen from the dynamics of the utility, profit and the consumption goods production levels. Even if in this specification the central bank directly buys government bonds from the government, which in principle has enough money to stimulate the economy, this instrument still cannot directly change already formed expectations, and therefore it has limited usefulness.

The results for production of capital goods at period t are given below:

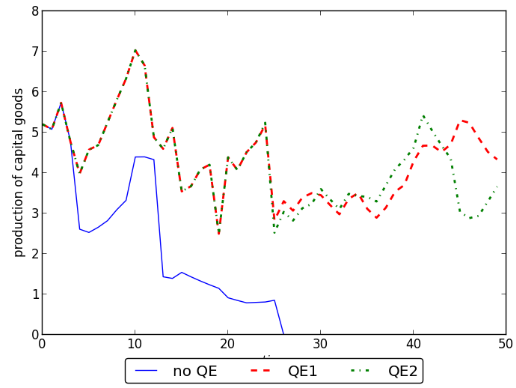

Figure 4.16 Simulation results for the capital good production levels, the case of simplified bankruptcy

The results for the production of consumption goods at period t are as follows:

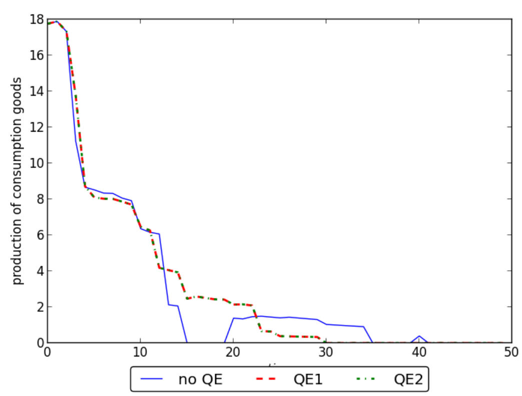

Figure 4.17 Simulation results for the consumption good production levels, the case of simplified bankruptcy

According to the results of the simulations, humans are not better off in the long term. It can be seen in the figure below, where the results for the average period utility calculated as in equation (A.21) are presented:

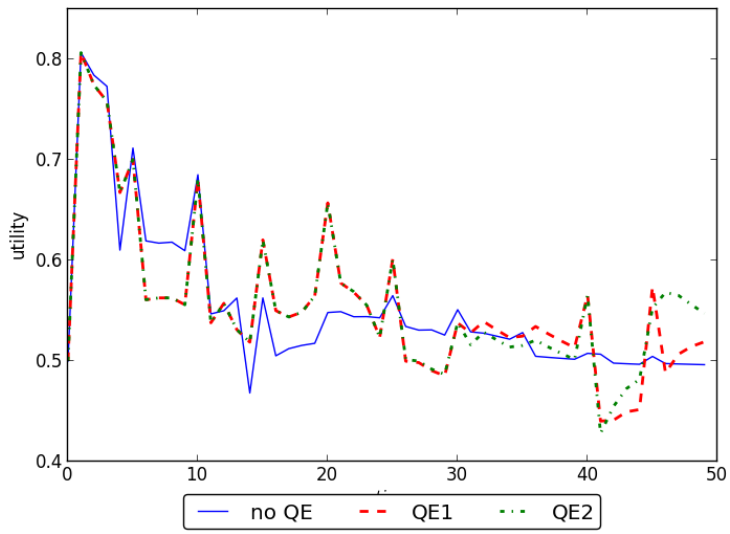

Figure 4.18 Simulation results for the average realized utility levels, the case of simplified bankruptcy

4.5.2.2 Results for the case of the bankruptcy with recapitalization arrangements

The following results for the main macroeconomic indicators refer to the case where bankruptcy framework is described by the *BankruptcyRestock* parameter. The levels of the credits, which are equal to the sum of all outstanding loans at time t in the economy, are presented below:

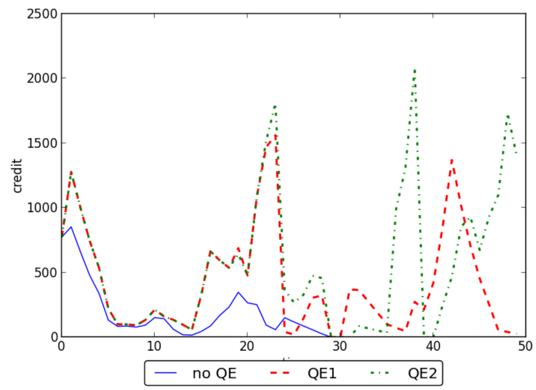

Figure 4.19 Simulation results for the credit levels, the case of bankruptcy with recapitalization

The agents in the economy take on excessive risks at the beginning, and eventually go bankrupt. In this economy QE policy exacerbate and support excessive risk taking which together with generous bankruptcy procedures leads to repeated boom-bust cycles in the economy.

The employment in period t is defined as a total number of active labor contracts. The employment levels are deteriorating in the simulations:

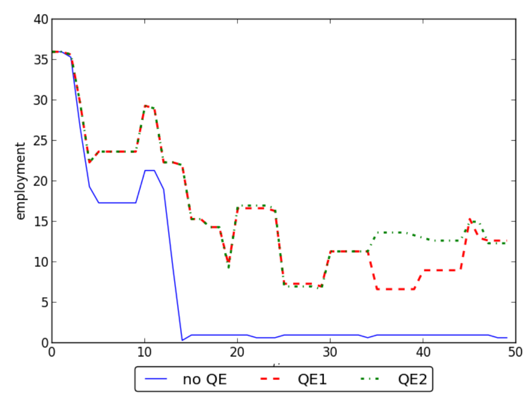

Figure 4.20 Simulation results for the employment levels, the case of bankruptcy with recapitalization

The level of deposits, defined by the total amount of outstanding deposits in the economy in period t, are presented below:

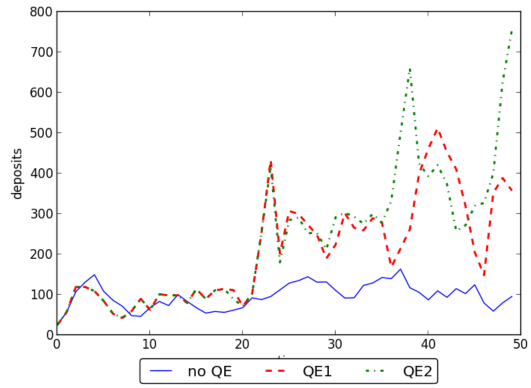

Figure 4.21 Simulation results for the deposit levels, the case of bankruptcy with recapitalization

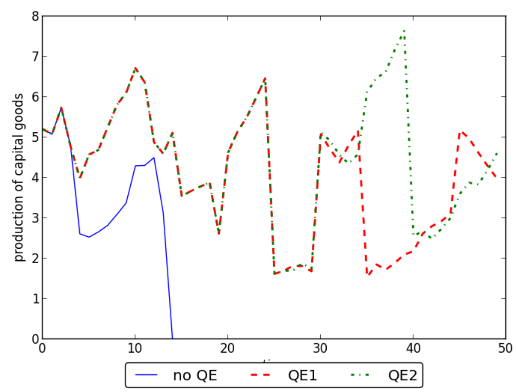

Figure 4.22 Simulation results for the capital good production levels, the case of bankruptcy with recapitalization

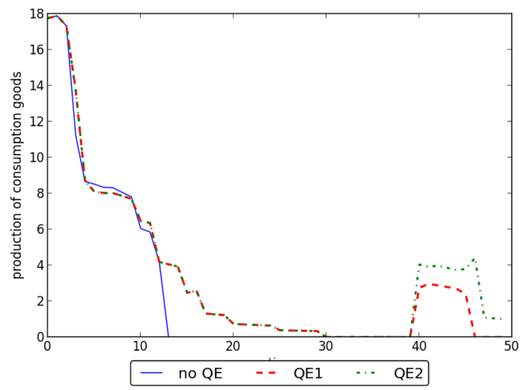

Figure 4.23 Simulation results for the consumption good production levels, the case of bankruptcy with recapitalization

By the end of the simulations, humans are once again not better off in the long term. The results are presented for the average realized utility calculated in the same way as in equation A.21

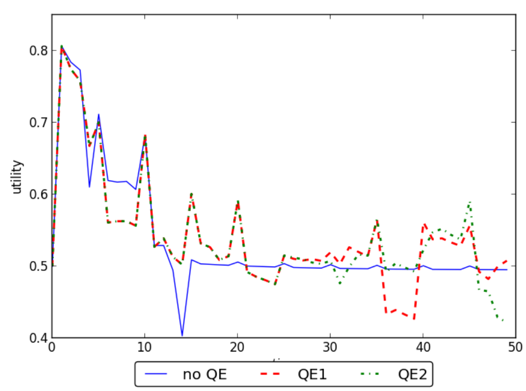

Figure 4.24 Simulation results for the average realized utility levels, the case of bankruptcy with recapitalization

The presented results differ from the case described in Section 4.5.2.1. Hence, regardless of the implemented policy, the outcome for the consumers and other agents depends on the initial structure of the economy, the forecasting rules and the institutional structure. It might be argued that to achieve major economic changes, the central bank instruments are insufficient, and complex policy decisions are required.

4.6 Conclusion

This chapter describes the results obtained from simulations with a relatively advanced model that includes not only traditional agents such as firms, consumers and the government, but also less frequently incorporated banks. This model was designed with the goal of creating a realistic simulation of the money and financial flows that would allow researchers to study network effects and effects of institutional regulations and incomplete information on the macroeconomic dynamics. This model is still being developed, but even now it is clear that the institutional regulations and restrictions play a significant role in the dynamics of the economy and in the choice of optimal policies.

The main conclusion that can be drawn from the simulations is that macroprudential policies should be evaluated in a model with heterogeneous agents and realistic institutional structures. A simplified model without any room for changes in the institutional structures and with perfectly shared information between all agents, such as the currently dominating DSGE models, is not able to capture significant effects of mistaken expectations, incomplete information on part of agents and effects arising from allowing too-big-to-fail institutions to exist.

In the short term, the central bank policies can influence the macroeconomic dynamics, but in the long term the existing institutional structure becomes the driving force behind the economic development. This result is consistent with the conclusions drawn from the previous chapters, where model structures were important diving forces behind particular outcomes.

CHAPTER 5. CONCLUSION

As a whole, the resented body of work explores the question of optimal behavioral choices made by agents and their effects on the micro- and the macro-level economic dynamics.

In the first chapter, the notion of constructive rationality is explained and applied to the analysis of behavioral choices made by agents. The decision process that satisfies the requirements of constructive rationality is introduced, and simulations are run that sample different specifications for the introduced decision rules. Various approximate algorithms that solve the posed optimization problem were explained and adapted to the problem on hand. These algorithms were tested further in the next chapter.

In the second chapter, another aspect of constructive rationality was explored. A number of possible expectation specifications for the economic dynamic was analyzed. It was concluded that the best choice of expectation formation rules heavily depends on particular characteristics of the environment (in the case of the simple model explored in that chapter, it was the degree of riskiness of the environment).

Finally, the last chapter presents a more advanced macroeconomic model. This model is designed to be a tool for the exploration of possible central bank policies. Consistently with the results of previous chapters, the effectiveness of the central bank policies is severely limited by the particular institutional framework characteristic of the economy.

There are many possible directions for future improvements. One would be to keep working on improving the modeling expectation formation and updating to make it closer to actual methods used by real agents. Another one would be to keep expanding the institutional framework until it becomes close enough to the real world, and then the resulting model will be able to provide insights into unintended interactions between different policies implemented simultaneously in the real world.

APPENDIX A. ADDITIONAL MATERIAL FOR CHAPTER 2

A.1 Tested Grid Specifications for Decision Domains

Table A.1 Small-grid discretization of the consumer decision domain D^c

Decision Component	Set of Possible Values
l^c	$L^c = \{0, 1\}$
ω	$\Omega = \{0.8, 1.0, 1.2\}$
θ	$\Theta = \{0.0, 0.5, 1.0\}$

Table A.2 Small-grid discretization of the firm decision domain D^f.

Decision Component	Set of Possible Values
l^f	$L^f = \{0, 2.5, 5.0, 7.5, 10\}$
γ	$\Gamma = \{0.8, 1.0, 1.2\}$
λ	$\Lambda = \{0.0, 0.5, 1.0\}$
ψ	$\Psi = \{0.8, 1.0, 1.2\}$

Table A.3 Big-grid discretization of the consumer decision domain D^c

Decision Component	Set of Possible Values
l^c	$L^c = \{0, 1\}$
ω	$\Omega = \{0.10, 0.55, 1.00, 1.45, 1.90\}$
θ	$\Theta = \{0.0, 0.5, 1.0\}$

Table A.4 Big-grid discretization of the firm decision domain D^f

Decision Component	Set of Possible Values
l^f	$L^f = \{0, 2.5, 5.0, 7.5, 10\}$
γ	$\Gamma = \{0.10, 0.55, 1.00, 1.45, 1.90\}$
λ	$\Lambda = \{0.0, 0.5, 1.0\}$
ψ	$\Psi = \{0.10, 0.55, 1.00, 1.45, 1.90\}$

A.2 Wage, Price, and Dividend Expectation Updating

Consumers and firms in the DM Game are assumed to follow the same methods in forming and updating their expectations regarding the distribution of future labor market wages, goods market prices, and dividend payments (for consumers). These methods are characterized by prior-belief parameters and a memory parameter. The prior-belief parameters are maintained parameters set at fixed values throughout all simulations reported in this study. The memory parameter is a treatment factor set to reflect either a fixed one-period memory or an expanding memory that takes into account all previous observations at each time t.

Let v denote any consumer or firm in the DM Game. At each time $t \geq 0$, agent v forms normal probability distributions for the labor market wage w, the goods market price p, and the dividend payment div in current and future periods. These normal probability distributions are characterized by state-conditioned estimates for their means and variances, as follows:

$$w \sim \mathcal{N}\left(\bar{w}_{v,t-1}, \sigma^{2\ L}_{v,t-1}\right) \tag{A.1}$$

$$p \sim \mathcal{N}\left(\bar{p}_{v,t-1}, \sigma^{2\ G}_{v,t-1}\right) \tag{A.2}$$

$$div \sim \mathcal{N}\left(\bar{d}_{v,t-1}, \sigma^{2\ D}_{v,t-1}\right) \tag{A.3}$$

After the determination of a market-clearing wage $w_{t:1}$, a market-clearing goods price $p_{t:3}$, and a dividend payment $div_{t:5}$ for period t, agent v updates the means and variances for these distributions in order to obtain updated estimates for these distributions for use in period $t+1$.

The method used to obtain updated mean and variance estimates for the wage distribution (A.1) is characterized by the following three parameters: a prior wage $w_{v,0}$; a prior weight $n^L_{v,0}$, and a memory parameter wm. If $wm = all$, then agent v calculates these estimates as follows:

$$\bar{w}_{v,t} = \frac{\sum_{r\ =\ 0}^{t} w_{r:1}\ +\ n^L_{v,0} \cdot w_{v,0}}{t\ +1\ +\ n^L_{v,0}} \tag{A.4}$$

$$\sigma^{2,L}_{v,t} = \frac{\sum_{r\ =\ 0}^{t} \left(w_{r:1}\ -\ \bar{w}_{v,t}\right)^2\ +\ n^L_{v,o} \cdot \left(w_{v,0}\ -\ \bar{w}_{v,t}\right)^2}{t\ +1\ +\ n^L_{v,0}} \tag{A.5}$$

In other words, the mean of the distribution for the expected wage is determined by averaging all wages observed to date, together with the prior wage, while the dispersion of the expected wage is determined by averaging the squares of the deviations of the observed wages and the prior wage from the currently estimated mean wage.

If $wm = one$, then agent v sets the expected wage equal to the most recently observed wage:

$$\bar{w}_{v,t} = w_{t:1} \tag{A.6}$$

Also, agent v sets the expected variance equal to 1% of this estimated expected wage:

$$\sigma^{2,L}_{v,t} = 0.01 \cdot \bar{w}_{v,t} \tag{A.7}$$

Similar equations are used to obtain updated estimates $\bar{p}_{v,t}$, $\sigma^{2,G}_{v,t}$, $\overline{div}_{v,t}$, and $\sigma^{2,D}_{v,t}$ for the means and variances for the goods price distribution (A.2) and the dividend distribution (A.3) for $wm = all$ and $wm = one$, with $p_{r:3}$ or $div_{r:5}$ replacing $w_{r:1}$, $p_{v,0}$ or $div_{v,0}$ replacing $w_{v,0}$, and $n^G_{v,0}$ or $n^D_{v,0}$ replacing $n^L_{v,0}$.

The estimated means $\bar{w}_{v,t}$ and $\bar{p}_{v,t}$ for the wage and the goods price are used to determine the reservation wage and reservation price for agent v's transformation function mapping described in Sections 2.3.2 and 2.3.3. Specifically, $E_{v,t}[w_{r:1}] = \bar{w}_{v,t-1}$ and $E_{v,t}[p_{r:3}] = \bar{p}_{v,t-1}$ for all $r \geq t$. Thus equations (2.14), (2.17), and (2.18) take the form

$$w^c_{i,r:1}(d,t) = \omega \cdot \bar{w}_{i,t-1} \tag{A.8}$$

$$w^f_{j,r:1}(d,t) = \gamma \cdot \bar{w}_{j,t-1} \tag{A.9}$$

$$p^f_{j,r:3}(d,t) = \lambda \cdot \bar{p}_{j,t-1} \tag{A.10}$$

As clarified below in Section A.3, the EO-FH and EO-ADP agents make use of the full probability distributions (A.1) through (A.3) in their decision processes. The updating of these distributions requires specifications for prior variance values as well as prior mean values.

A complete listing of the maintained values for all of the prior-belief parameters is given in Table A.5.

Table A.5 Maintained values for prior-belief parameters

Parameter	Value
$w_{v,0}$	1.00
$p_{v,0}$	1.00
$div_{v,0}$	0.00
$n^{L}_{v,0}$	10.00
$n^{G}_{v,0}$	10.00
$n^{D}_{v,0}$	0.00
$\sigma^{2\ L}_{v,0}$	0.50
$\sigma^{2\ G}_{v,0}$	0.50
$\sigma^{2\ D}_{v,0}$	0.01

A.3 Implementation of EO Decision Rules

Various computational approximation methods could be used to implement the EO-FH and EO-ADP decision procedures. The methods used in this study are outlined below. Detailed explanations of these methods can be found in Powell (2011).

A.3.1 Implementation of the EO-ADP Decision Rule

By assumption, consumers in the DM Game are structurally identical. In particular, they have the same form of budget and feasibility constraints (2.1) through (2.4), the same intertemporal utility objective function (2.5), and the same single-period utility function $u(\cdot)$ given by (2.37).

The state $\mathbf{x}_{i,t}$ of any consumer i at any time $t \geq 0$ is given by:

$$\mathbf{x}_{i,t} \;=\; \left[t, M^{c}_{i,t-1}, \bar{w}_{i,t-1}, \sigma^{2\ L}_{i,t-1}, \bar{p}_{i,t-1}, \sigma^{2\ G}_{i,t-1}, \bar{div}_{i,t-1}, \sigma^{2\ div}_{i,t-1}\right] \tag{A.11}$$

The dimension of the state (A.11) is fixed at eight, independently of i and t. Our normality assumptions imposed on the wage, price, and dividend payment distributions (A.1) through (A.3) implies that each of these distributions is fully characterized in each period t by its estimated mean and variance appearing in (A.11).

The value function for consumer i at time t in state $\mathbf{x}_{i,t}$ takes the form:

$$V^c\left(\mathbf{x}_{i,t}\right) \;=\; \max_{d \in D^c} E_{i,t} \sum_{r=t}^{\infty} \beta^{r-t} \left[u\left(q^c_{i,r:3}(p_{r:3}, d, t), 1 - l^c_{i,r:1}(w_{r:1}, d, t)\right)\right] \qquad \text{(A.12)}$$

subject to: the budget and feasibility constraints (2.1) through (2.4) that depend on $\mathbf{w}_{t:1}$, $\mathbf{p}_{t:3}$, and $\mathbf{div}_{t:5}$; and (ii) the $\mathbf{TR}^c_{i,t}$ function that maps each potential period-t decision $d \in D^c$ into a sequence of labor supply and goods demand functions for periods $r \geq t$. The expectation in (A.12) is taken with respect to the wage, price, and dividend payment probability distributions (A.1) through (A.3), conditional on $\mathbf{x}_{i,t}$.

The state transition function $\mathbf{S}^c$ mapping each possible state $\mathbf{x}_{i,t}$, decision $d \in D^c$, and realization $(w_{t:1}, p_{t:3}, div_{t:5})$ for the wage, price, and dividend payment in period t into an updated state $\mathbf{x}_{i,t+1}$ for period $t+1$ is time invariant and the same for all consumers i. Also, the left-side summation in (A.12) is time separable. Consequently, the value function $V^c(x_{i,t})$ can equivalently be expressed in recursive form, as follows:

$$\begin{aligned} V^c\left(\mathbf{x}_{i,t}\right) = \max\nolimits_{d \in D^c} E_{i,t} &\left[u\left(q^c_{i,t:3}(p_{t:3}, d, t), 1 - l^c_{i,t:1}(w_{t:1}, d, t)\right)\right. \\ &+ \left.\beta V^c\left(\mathbf{S}^c\left(\mathbf{x}_{i,t}, d, w_{t:1}, p_{t:3}, div_{t:5}\right)\right)\right] \end{aligned} \qquad \text{(A.13)}$$

We assume that each EO-ADP consumer i at each time t derives an estimate for the value function (A.12) that solves the recursive relationship (A.13) by means of a type of *adaptive dynamic programming (ADP)* algorithm surveyed in (Powell, 2011, p. 407). The latter algorithm, designed for infinite-horizon dynamic programming problems, is an approximate policy iteration method implemented by means of least-squares temporal differencing.

During this value function estimation at time t, the mean and variance estimates $\bar{w}_{i,t-1}$, $\sigma^{2\,L}_{i,t-1}$, $\bar{p}_{i,t-1}$, $\sigma^{2\,G}_{i,t-1}$, $\bar{d}_{i,t-1}$, and $\sigma^{2\,div}_{i,t-1}$ in consumer i's state $\mathbf{x}_{i,t}$ are held fixed. No new information is obtained by consumer i during his value function estimation, so he does not update his state information during this estimation.

A critical step in the EO-ADP algorithm at each time t is the selection of basis functions for approximating the general form of the value function prior to conducting the value function estimation. We assume each EO-ADP consumer i at each time t uses a single linear basis function, as follows:

$$V^c(\mathbf{x}_{i,t}) \;=\; \sum_k \theta_k^{\pi} \phi_k\left(\mathbf{x}_{i,t}\right) \;=\; \theta^{\pi} \cdot M^c_{i,t-1} \tag{A.14}$$

where $M^c_{i,t-1}$ denotes the time-t money balance of consumer i. The value function estimation problem at time t thus reduces to the estimation of the scalar parameter θ^{π} over some specified domain, which in this study was taken to be the interval $[0.01, 1000]$.

It is assumed that EO-ADP firms use a similar EO-ADP decision procedure to estimate their time-t value functions. The state $\mathbf{x}_{i,t}$ of an EO-ADP firm j at time t is given by

$$\mathbf{x}_{j,t} \;=\; \left(t, M^f_{i,t-1}, \bar{w}_{j,t-1}, \sigma^{2\ L}_{j,t-1}, \bar{p}_{j,t-1}, \sigma^{2\ G}_{j,t-1}\right) \tag{A.15}$$

and its value function is given by

$$V_t^f\left(\mathbf{x}_{j,t}\right) \;=\; \max_{d \in D^f} E_{j,t} \sum_{r=t}^{\infty} \mu^{r-t} \left[p_{r:3} q^f_{j,r:3}(p_{r:3}, d, t) - w_{r:1} l^f_{j,r:1}(w_{r:1}, d, t)\right] \tag{A.16}$$

The right-side maximization in (A.16) is constrained by the technological and feasibility constraints (2.6) through (2.11), conditional on $\mathbf{x}_{j,t}$, and implicitly depends on the $\mathbf{TR}^f_{j,t}$ function that maps each potential period-t decision $d \in D^f$ into a sequence of labor demand and goods supply functions for periods $r \geq t$. The expectation in (A.16) is taken with respect to the wage and price probability distributions (A.1) and (A.2), conditional on $\mathbf{x}_{j,t}$.

For reasons analogous to arguments given above for EO-ADP consumers, the value function (A.16) can be expressed in the following recursive form:

$$\begin{aligned} V^f\left(\mathbf{x}_{j,t}\right) \;=\; & \max_{d \in D^f} E_{j,t} \Big[p_{t:3} q^f_{j,t:3}(p_{t:3}, d, t) - w_{t:1} l^f_{j,t:1}(w_{t:1}, d, t) \\ & + \;\; \beta V^f\left(\mathbf{S}^f\left(\mathbf{x}_{j,t}, d, w_{t:1}, p_{t:3}\right)\right)\Big] \end{aligned} \tag{A.17}$$

where the form $\mathbf{S}^f$ of the state transition function does not depend on j or t. Firm j at time t is assumed to use a simple linear basis function to estimate the value function $V^f(\mathbf{x}_{j,t})$ that solves (A.17), as follows:

$$V^f(\mathbf{x}_{j,t}) \;=\; \sum_z \theta_z^\pi \phi_z\left(\mathbf{x}_{j,t}\right) \;=\; \theta^\pi \cdot M_{j,t-1}^f \tag{A.18}$$

where $M_{j,t-1}^f$ denotes the money balance of firm j at time t.

The following parameters need to be specified in order to implement the EO-ADP algorithm for EO-ADP consumers and EO-ADP firms: the number of runs for the inside and outside estimation loops; the number of random number draws in an internal maximization algorithm ; the number of basis functions; the initial parameter value B^0 for recursive least squares estimation; and the initial parameter value $\theta^{\pi,0}$ for the coefficient in the basis-function representation of the value function. These parameters are maintained at the fixed values listed in Table A.6 for all EO-ADP agents. The tested values for the two EO-ADP treatment factors, wm and grid-type, are given in Table 2.10.

Table A.6 Maintained parameter values for EO-ADP agents

Parameter	Value
EstRunIn	5
EstRunOut	5
BasisNum	1
NDrawsADP	5
B^0	$0.005 \cdot I$
$\theta^{\pi,0}$	1.0

A.3.2 Implementation of the EO-FH decision rule

The EO-FH algorithm is a brute-force method for the direct estimation of an optimal solution in each period t over a finite rolling forecasting-horizon T. It is performed by

EO-FH consumers and firms by undertaking a complete search of their finite decision domains, with a corresponding evaluation of expected outcomes over the next T periods, in order to determine a decision achieving the maximum possible expected intertemporal utility or profit outcome over these next T periods. Thus, in contrast to the EO-ADP algorithm, the EO-FH algorithm does not involve estimation over an infinite horizon, and it does not involve the use of value functions. Consequently, it is conceptually simpler and faster to implement than the EO-ADP algorithm.

Specifically, each EO-FH consumer i at each time t in some state $\mathbf{x}_{i,t}$ uses direct search to solve an optimization problem identical in form to (A.12) except that the infinite horizon is replace by a finite horizon $t+T$. Similarly, each EO-FH firm j at each time t in some state $\mathbf{x}_{j,t}$ uses direct search to solve an optimization problem identical in form to (A.16) except that the infinite horizon is replace by a finite horizon $t+T$.

The EO-FH consumers and firms at each time t use Monte Carlo simulation to calculate the expectations in their finite-horizon maximization problems, by taking NDrawsFH draws from each of their estimated probability distributions (A.1), (A.2), and (A.3). The value of the parameter NDrawsFH is maintained at 10 for all EO-FH agents. The tested values for the three EO-FH treatment factors T, wm, and grid-type are given in Table 2.9.

A.4 Social Planner Benchmark Model Solution

This section provides a proof by induction that the Social Planner (SP) Benchmark Model in reduced representative-consumer form (2.41) has the following solution: $l^c_{t:1} = q^c_{t:3} = 1$ and $s^{stock}_t = 0$ for all periods $t \geq 0$.

By assumption, $s^{stock}_{-1} = 0$. Given this assumption, the social planner's optimal choices for labor, consumption, and goods stock for period 0 are given by $l^c_{0:1} = q^c_{0:3} = 1$ and $s^{stock}_0 = 0$. To establish this, first note that leisure $le^c_{0:1} = [1 - l^c_{0:1}]$ has a constant marginal utility equal to 0.5 whereas goods consumption $q^c_{0:3}$ over the range $(0, 1]$ has a marginal utility that is bounded below by 1.5. Consequently, the social planner will set $le^c_{0:1} = 0$ (hence $l^c_{0:1} = 1$). Given the production function assumptions for the SP Benchmark Model, the maximum amount of good that can be produced in period 0 is thus 1 unit.

Now suppose the social planner contemplates setting aside a portion $s_0^{stock} \in [0, 1]$ of this period-0 production as goods stock for period 1. Given s_0^{stock}, the maximum utility achievable in period 0 by the representative consumer is $3.0 \ln(2 - s_0^{stock})$ if $s_0^{stock} < 1$ and $3.0 \ln(0.5)$ if $s_0^{stock} = 1$. Also, given s_0^{stock}, the maximum utility achievable by the representative consumer in period 1 is then attained by setting $l^c_{r:1} = 1$, allocating all of the resulting period-1 production of 1 unit of good towards time-1:3 consumption, and allocating all of the goods stock s_0^{stock} towards time-1:3 consumption,. From the standpoint of period 0, the resulting maximum utility achievable by the representative consumer in period 1 is thus given by $\beta[3.0 \ln(2 + s_0^{stock})]$. However, since β is less than 1, the sum of these two maximum achievable utility levels,

$$3.0 \ln(2 - s_0^{stock}) + \beta \cdot \left[3.0 \ln(2 + s_0^{stock})\right] \quad , \tag{A.19}$$

is a strictly decreasing function of s_0^{stock} over $s_0^{stock} \in [0, 1)$ (with a discontinuous further jump down at $s_0^{stock} = 1$). Consequently, the maximum achievable intertemporal utility for the representative consumer over periods 0 and 1, considered together, is obtained by setting $s_0^{stock} = 0$. Similar arguments can be used to argue that no future use of a positive s_0^{stock} can result in a (discounted) utility gain for the representative consumer that outweighs his resulting loss of period-0 utility. Consequently, the social planner should set $s_0^{stock} = 0$.

Now consider any arbitrary period $t \geq 0$ for which the goods stock s_{t-1}^{stock} is zero. Then the same argument used above can be applied to period t to show that the social planner's optimal choices for period t are to set $l^c_{t:1} = q^c_{t:3} = 1$ and $s_t^{stock} = 0$. It follows by induction that the optimal solution to the SP Benchmark Model (2.41) is $l^c_{t:1} = q^c_{t:3} = 1$ and $s_t^{stock} = 0$ for all periods $t \geq 0$.

A.5 Performance Measures for Case Comparisons

Let k denote any of the tested cases in Table 2.3. This section describes the various performance measures used to evaluate the performance of the DM-Game economy under case k.

The primary indicator used to measure ex post performance is $\bar{u}^k$, the *average realized single-period utility* attained by the DM-Game consumers. Using notation introduced in Section 2.5.1, and recalling that the initial period is numbered 0, $\bar{u}^k$ is calculated as follows:

$$\bar{u}^k \;=\; \frac{\sum_{i=1}^{I}\sum_{\tau=LOmit}^{LRun}\sum_{r=1}^{NRuns} u^k_{i,\tau,r}}{I\cdot(LRun - LOmit + 1)\cdot NRuns} \tag{A.20}$$

where $u^k_{i,\tau,r}$ is the utility attained by consumer i in period τ of run r.

Some use is also made of additional performance measures. For each period $\tau \in \{\text{LOmit},\ldots,\text{LRun}\}$, the *average realized single-period utility for period* τ is calculated as follows:

$$\bar{u}^k_\tau \;=\; \frac{\sum_{i=1}^{I}\sum_{r=1}^{NRuns} u^k_{i,\tau,r}}{I\cdot NRuns} \tag{A.21}$$

The average value of $\bar{u}^k_\tau$ across the time periods $\tau \in \{\text{LOmit},\ldots,\text{LRun}\}$ is then given by (A.20), and the standard deviation of $\bar{u}^k_\tau$ across these same time periods is given by

$$\sigma_{\bar{u}^k} \;=\; \left(\frac{\sum_{\tau=LOmit}^{LRun}\left(\bar{u}^k_\tau \;-\; \bar{u}^k\right)^2}{LRun - LOmit + 1}\right)^{1/2} \tag{A.22}$$

The *average realized cumulative utility through period* t is calculated as follows for periods $t \geq \text{LOmit}$:

$$\bar{u}^{cumul,k}_t \;=\; \frac{\sum_{\tau=LOmit}^{t}\bar{u}^k_\tau}{t - \text{LOmit} + 1} \tag{A.23}$$

Suppose that a market-clearing wage $w^k_{t:1,r}$ and a market-clearing goods price $p^k_{t:3,r}$ are both well-defined[1] for some period t for all runs $r \in R^*$, where the subset R^* has

[1] Since the demands and supplies of the DM-Game consumers and firms depend on reservation wages and prices, there can exist periods for which all of these agents decide not to participate in the labor market and/or the goods market.

cardinality $NRuns^*$. Then the *average realized real wage for period* t is calculated as follows:

$$\bar{w}_t^{real,k} \;=\; \frac{\sum_{r=1}^{NRuns^*}\left[\frac{w_{t:1,r}^k}{p_{t:3,r}^k}\right]}{NRuns^*} \tag{A.24}$$

The *average realized real wage* $\bar{w}^{real,k}$ is then calculated as the average of $\bar{w}_t^{real,k}$ over all periods t for which $\bar{w}_t^{real,k}$ is well defined.

Finally, in analogy to (A.20), the *average realized single-period profits* attained by the DM-Game firms is calculated as follows:

$$\bar{\pi}^k \;=\; \frac{\sum_{j=1}^{J}\sum_{\tau=LOmit}^{LRun}\sum_{r=1}^{NRuns}\pi_{j,\tau,r}^k}{J\cdot(LRun-LOmit+1)\cdot NRuns} \tag{A.25}$$

where $\pi_{j,\tau,r}^k$ denotes the profit attained by firm j in period τ of run r.

APPENDIX B. ADDITIONAL MATERIAL FOR CHAPTER 3

B.1 Initial Parameters for Chapter 3 Simulations

The initial parameters for Chapter 3 are given in Table B.1.

Table B.1 Other simulation parameters for Chapter 3

Parameter	**Value**
β	0.95
$wealth$	1.0
$lifespan\ T$	3
$inheritance$	$full$
ADP_N	10
ADP_M	10
CS_N	100
$CS\_N\_discret$	10
$LRun$	99 or 999 for 1 level prior
B^0	$0.0005 \cdot I$
$\theta^{\pi,0}$	$[100, 100, 100]$
$seed$	$[2012, 2013, 2014]$
$NSeed$	3

B.2 Approximate Dynamic Programming Algorithm

B.2.1 Core ADP algorithm

At each step, new estimates for the value function are obtained. After that, coefficients for the value function linearization are updated using the least squares algorithm

2 presented below. This process repeats for a fixed number of steps. The main body of the algorithm 1 is given below.

Algorithm 1 ADP algorithm for estimation of value functions.

Step 0 Initialization
Step 0a Fix the basis functions $\phi_f(s)$
Step 0b. Initialize $\theta_{tf}^{\pi,0}$
Step 0c. Set $n = 1$
Step 1. Sample an initial starting state X_0^n:
Step 2. Initialize $\theta^{n,0}$(if $n > 1$, use $\theta^{n,0} = \theta^{n-1}$), which is used to estimate the value of policy π produced by $\theta^{\pi,n}$. $\theta^{n,0}$ is used to approximate the values of the following policy π determined by $\theta^{\pi,n}$
Step 3. Do for $m = 1, 2, \ldots, M$:
Step 4. Choose a sample path ω^m.
Step 5. Do (Steps 5a, 5b)
Step 5a. Compute d (using grid search, see below)

$$d_t = \underset{d_t \in D_t^{n,m}}{\arg\max} \left(C\left(X_t^{n,m}, d_t\right) + \gamma \mathbb{E} \left(\sum_f \theta_{tf}^{\pi,n-1} \phi_f \left(TR\left(X_t^{n,m}, d_t\right)\right) \right) \right)$$

Step 5b. Compute

$$X_{t+1}^{n,m} = TR\left(X_t^{n,m}, d, W_{t+1}\left(\omega^m\right)\right)$$

Step 6. Initialize

$$v_{T+1}^{n,m} = 0$$

Step 6. For t=T,T-1,...,0

$$v_t^{n,m} = C\left(X_t^{n,m}, d_t^{n,m}\right) + \gamma \mathbb{E}\left(v_{t+1}^{n,m}\right)$$

Step 7. Update $\theta^{n,m-1}$ using recursive least squares to obtain $\theta^{n,m}$, go to Step 3
Step 8. Set $n = n + 1$. If $n < N$, go to step 1.
Step 9. Return the regression coefficients θ^N.

To actually solve step 5.a of the ADP algorithm, a grid search is used, with the same assumptions as in the complete grid search approach described below.

B.2.2 Complete Grid Search

The complete grid search discretizes possible choices x_t over the interval of $[0, 1]$. After that, the complete search over all possible choices for one period in case of ADP

Algorithm 2 Recursive least squares for ADP algorithm.

Step 0. If $n = 1$, initialize $B^0 = \epsilon I$. Else use B^{n-1}.
Step 1. Calculate error

$$\epsilon^n = \bar{V}_s\left(\theta^{n-1}\right) - v^n$$

Step 2. Calculate adjustment coefficient γ^n.

$$\gamma^n = 1 + \left(\phi^n\right)' B^{n-1}\phi^n$$

Step 3. Calculate B^n.

$$B^n = B^{n-1} - \frac{1}{\gamma^n}\left(B^{n-1}\phi^n\left(\phi^n\right)' B^{n-1}\right)$$

Step 4. Calculate H^n.

$$H^n = \frac{1}{\gamma^n}B^{n-1}$$

Step 5. Calculate new regression coefficient estimates θ^n and store B^n.

$$\theta^n = \theta^{n-1} - H^n\phi^n\epsilon^n$$

or T periods of the total life span length for the benchmark algorithm are run. The best choice combination is chosen.

APPENDIX C. ADDITIONAL MATERIAL FOR CHAPTER 4

Initial Parameters for Chapter 4 Simulations

The initial parameters for Chapter 4 are given below.

Table C.1 Main parameter values

Parameter	**Description**	**Value**
nH	number of humans	100
$nFGC$	number of consumption good firms	2
$nFGK$	number of capital goods firms	2
nB	number of banks	2
nCB	number of central banks	1
nG	number of governments	1
β	discounting factor	0.95
$fgc\_F\_F\_theta$	production function parameters for FGC	$(1, 0.7, 0.3)$
$fgk\_F\_F\_theta$	production function parameters for FGK	$(0.3, 0.3, 0.7)$
$b\_F\_F\_min$	production function parameters for banks	$(1.0, 1.0)$
$h\_goal\_t\_theta$	utility function parameters for humans	$(1.0, 3.0, 0.5)$
wm	memory length	*all*
$bankruptcy\_length$	length of a bankruptcy	4

Table C.2 Parameter values for simulations

Parameter	**Description**	**Value**
N	number of periods in a simulation	50
$NRuns$	number of seeds per simulation	3
$seeds$	seeds for simulation	$[2014, 104, 255]$
ADP_M	M value for ADP algorithm	5
ADP_N	N value for ADP algorithm	5
CS_N	N value for complete search algorithm	5

Table C.3 Discretization of the human decision domain D^h

Decision Component	**Set of Possible Values**
θ^L	$\Theta^L = \{1\}$
ω^L	$\Omega^L = \{0.5, 1.0, 1.5\}$
θ^K	$\Theta^K = \{0.0, 0.3\}$
θ^C	$\Theta^C = \{0.3\}$
θ^{Cr}	$\Theta^{Cr} = \{1.0, 1.5\}$
θ^{Dp}	$\Theta^{Cr} = \{0.3\}$

Table C.4 Discretization of the firm decision domain D^f

Decision Component	**Set of Possible Values**
θ^L	$\Theta^L = \{1.0\}$
ω^L	$\Omega^L = \{0.5, 1.0, 1.5\}$
θ^K	$\Theta^K = \{1.0\}$
θ^G	$\Theta^G = \{1.0\}$
ω^G	$\Omega^C = \{1.0\}$
θ^{Cr}	$\Theta^{Cr} = \{0.0, 0.5, 1.0\}$
θ^{Dp}	$\Theta^{Cr} = \{0.1\}$

Table C.5 Discretization of the bank decision domain D^b

Decision Component	**Set of Possible Values**
θ^{Cr}	$\Theta^{Cr} = \{10.0\}$
θ^{Cr}	$\theta^{Cr} = \{0.05, 0.1, 0.15, 0.2\}$
θ^{Dp}	$\theta^{Dp} = \{0.05, 0.1, 0.15, 0.2\}$
θ^{Bg}	$\Theta^{Bg} = \{0.1\}$
θ^{Bb}	$\Theta^{Bb} = \{0.05, 0.1, 0.15, 0.2\}$

Table C.6 Parameter values for expectation updating beg.

Parameter	Description	Value
MasCHK_mu	market for labor, initial mean	1.0
MasCHK_n0	market for labor, initial number of observations	10
MasCHK_sigma2	market for labor, initial variance	1.0
MasGK_mu	market for capital, initial mean	10.0
MasGK_n0	market for capital, initial number of observations	10
MasGK_sigma2	market for capital, initial variance	1.0
MasGC_mu	market for goods, initial mean	1.0
MasGC_n0	market for goods, initial number of observations	10
MasGC_sigma2	market for goods, initial variance	1.0
MasCBC_mu	market for credit, initial mean	0.1
MasCBC_n0	market for credit, initial number of observations	10
MasCBC_sigma2	market for credit, initial variance	0.05
MasCBD_mu	market for deposits, initial mean	0.07
MasCBD_n0	market for deposits, initial number of observations	10
MasCBD_sigma2	market for deposits, initial variance	0.05
MasCBD_mu	market for deposits, initial mean	0.07
MasCBD_n0	market for deposits, initial number of observations	10
MasCBD_sigma2	market for deposits, initial variance	0.05

Table C.7 Parameter values for expectation updating cont.

Parameter	**Description**	**Value**
MasCBB_mu	interbank market, initial mean	0.05
MasCBB_n0	interbank market, initial number of observations	10
MasCBB_sigma2	interbank market, initial variance	0.05
MasBG_mu	market for bonds, initial mean	10.0
MasBG_n0	market for bonds, initial number of observations	10
MasBG_sigma2	market for bonds, initial variance	0.0
MasCCBC_mu	standing facilities of central bank, initial mean	0.1
MasCCBC_n0	standing facilities of central bank, initial number of observations	10
MasCCBC_sigma2	standing facilities of central bank, initial variance	0.0
ai_mu	average income for human, initial mean	1.0
ai_n0	average income for human, initial number of observations	10
ai_sigma2	average income for human, initial variance	0.0
FI_I_mu	financial income for human, initial mean	0.0
FI_I_n0	financial income for human, initial number of observations	10
FI_I_sigma2	financial income for human, initial variance	0.0

Table C.8 Parameter values for decisions

Parameter	Description	Value
$\theta^b_{c^h}$	bank choice of the share of income to be used in lending	1.0
$\theta^b_{c^f}$	bank choice of the share of assets to be used in lending	1.0
$share\_out$	estimated share of payments that go through other banks	0.5
$share\_ass$	estimated share of deposits to receive from other banks	0.5
$\theta^{f,div}$	share of the firm net profit to pay as dividends	0.5
$\theta^{b,div}$	share of the bank net profit to pay as dividends	0.5
$\theta^{cb,div}$	share of the central bank net profit to pay as dividends	1.0
ω_{Cb}	interest rate on standing facilities of the central bank	0.01

Table C.9 Parameter values for markets

Parameter	**Description**	**Value**
c_length_Hk	the length of a labor contract	5
c_length_Cr	the length of a credit contract	5
c_length_Dp	the length of a deposit contract	5
c_length_Bg	the length of a bond	6
c_length_Bb	the length of an interbank loan contract	1
$\theta_{min,pG}$	minimum coefficient for the price for the goods market bids formation	1
$\theta_{max,pG}$	maximum coefficient for the price for the goods market bids formation	1

Table C.10 Parameter values for initial stocks

Parameter	**Description**	**Value**
$M_{g,-1}$	initial money balances for government	5
$M_{cb,-1}$	initial money balances for central bank	0
$M_{f,-1}$	initial money balances for firms	100.0
$M_{b,-1}$	initial money balances for banks	100.0
$M_{h,-1}$	initial money balances for humans	1.0
$K_{b,-1}$	initial capital for banks	10.0
$K_{f,-1}$	initial capital for firms	10.0
$K_{h,-1}$	initial capital for humans	0.0
$p_{stock,b,-1}$	price of a stock for initial tier 1 capital for banks	1.0
$q_{stock,b,-1}$	quantity of a stock for initial tier 1 capital for banks	nH

BIBLIOGRAPHY

Ahn, D., Choi, S., Gale, D., and Kariv, S. (2007). Estimating ambiguity aversion in a portfolio choice experiment.

Alden, J. M. and Smith, R. L. (1992). Rolling horizon procedures in nonhomogeneous markov decision processes. *Operations Research*, 40(3):S183–S194.

Ashraf, Q., Gershman, B., and Howitt, P. (2011). Banks, market organization, and macroeconomic performance: an agent-based computational analysis. Technical report, National Bureau of Economic Research.

Bach, D., Seymour, B., and Dolan, R. (2009). Neural activity associated with the passive prediction of ambiguity and risk for aversive events. *The Journal of Neuroscience*, 29(6):1648.

Bernanke, B. S. and Reinhart, V. R. (2004). Conducting monetary policy at very low short-term interest rates. *American Economic Review*, pages 85–90.

Borghans, L., Heckman, J., Golsteyn, B., and Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3):649–658.

Carbone, E. and Hey, J. (2004). The effect of unemployment on consumption: an experimental analysis. *The Economic Journal*, 114(497):660–683.

Chen, S.-H. (2012). Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective. *Journal of Economic Dynamics and Control*, 36(1):1–25.

Christiano, L, M. R. e. M. R. (2010). Financial factors in economic fluctuations. Technical report, ECB working paper, 1192.

Curdia, V. and Woodford, M. (2010). Conventional and unconventional monetary policy. *Federal Reserve Bank of St. Louis Review*, 92(July/August 2010).

Curdia, V. and Woodford, M. (2011). The central-bank balance sheet as an instrument of monetary policy. *Journal of Monetary Economics*, 58(1):54 – 79.

Dawid, H., Gemkow, S., Harting, P., van der Hoog, S., and Neugart, M. (2011). The eurace@ unibi model: An agent-based macroeconomic model for economic policy analysis. *Report, Bielefeld University.*

De Resende, C., Dib, A., Lalonde, R., and Perevalov, N. (2013). Countercyclical bank capital requirement and optimized monetary policy rules. Technical report, Bank of Canada Working Paper.

Dosi, G. and Egidi, M. (1991). Substantive and procedural rationality. *Journal of Evolutionary Economics*, 1:145–168.

Dosi, G., Fagiolo, G., Napoletano, M., and Roventini, A. (2013). Income distribution, credit and fiscal policies in an agent-based keynesian model. *Journal of Economic Dynamics and Control*, 37(8):1598–1625.

Dosi, G., Fagiolo, G., and Roventini, A. (2010). Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767.

Duffy, J. (2008). Macroeconomics: a survey of laboratory research. *Handbook of Experimental Economics*, 2.

Epstein, L. and Schneider, M. (2010). Ambiguity and asset markets. Technical report, National Bureau of Economic Research.

Erev, I. and Roth, A. E. (1998). Predicting how people play games with unique, mixed-strategy equilibria. *American Economic Review*, 88:848–881.

Galati, G. and Moessner, R. (2013). Macroprudential policy - a literature review. *Journal of Economic Surveys*, 27(5):846–878.

Gertler, M. and Karadi, P. (2011). A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17 – 34. Carnegie-Rochester Conference Series on Public Policy: The Future of Central Banking April 16-17, 2010.

Gertler, M. and Karadi, P. (2013). Qe 1 vs. 2 vs. 3...: A framework for analyzing large-scale asset purchases as a monetary policy tool. *International Journal of Central Banking*, 9(1):5–53.

Gertler, M. and Kiyotaki, N. (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of monetary economics*, 3(11):547–599.

Gode, D. K. and Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy*, 101:119–137.

Hayashi, T. and Miao, J. (2010). Intertemporal substitution and recursive smooth ambiguity preferences. *Theoretical Economics, forthcoming.*

Hey, J. and Knoll, J. (2011). Strategies in dynamic decision making-an experimental investigation of the rationality of decision behaviour. *Journal of Economic Psychology.*

Ho, J., Robin Keller, L., and Keltyka, P. (2002). Effects of outcome and probabilistic ambiguity on managerial choices. *Journal of Risk and Uncertainty*, 24(1):47–74.

Howitt, P. (2008). Macroeconomics with intelligent autonomous agents. In Farmer, R., editor, *Macroeconomics in the Small and the Large: Essays on Microfoundations, Macroeconomic Applications and Economic History in Honor of Axel Leijonhufvud*. Edward Elgar.

Howitt, P. (2012). What have central bankers learned from modern macroecononomic theory? *Journal of Macroeconomics*, 34(1):11–22.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, 310(5754):1680.

Joyce, M., Miles, D., Scott, A., and Vayanos, D. (2012). Quantitative easing and unconventional monetary policy - an introduction. *The Economic Journal*, 122(564):F271–F288.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Keck, S., Diecidue, E., and Budescu, D. (2010). Group decision making under vagueness.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, 73(6):1849–1892.

Maffioletti, A., Schmidt, U., and Schroder, C. (2009). The effect of elicitation methods on ambiguity aversion: an experimental investigation. *Economics Bulletin*, 29(2):638–643.

Mandel, A., Jaeger, C., Fürst, S., Lass, W., Lincke, D., Meissner, F., Pablo-Marti, F., Wolf, S., et al. (2010). Agent-based dynamics in disaggregated growth models. *Université Paris1 Panthéon-Sorbonne (Post-Print and Working Papers)*.

Nicolaisen, J., Petrov, V., and Tesfatsion, L. (2001). Market power and efficiency in a computational electricity market with discriminatory double-auction pricing. *IEEE Transactions on Evolutionary Computation*, 5(5):504–523.

Oeffner, M. (2008). Agent-based keynesian macroeconomics - an evolutionary model embedded in an agent-based computer simulation. *Dissertation, Julius-Maximilians-Universitaet Wuerzburg, Germany.*

Powell, W. (2014). Clearing the Jungle of Stochastic Optimization: INFORMS Tutorial.

Powell, W. B. (2011). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, volume 842. Wiley.

Roth, A. E. and Erev, I. (1995). Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8:164–212.

Sbordone, A. M., Tambalotti, A., Rao, K., and Walsh, K. (2010). Policy analysis using DSGE models: An introduction. *FRBNY Economic Policy Review.*

Shefrin, H. M. and Thaler, R. H. (1988). The behavioral life-cycle hypothesis. *Economic Inquiry*, 26(4):609–643.

Simon, H. A. (1978). Rationality as process and as product. *American Economic Reviewr*, 68(2):1–12.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175.

Smith, V. L. (2008). *Rationality in Economics: Constructivist and Ecological Forms.* Cambridge University Press.

Stiglitz, J. E. (2002). Information and the change in the paradigm in economics. *The American Economic Review*, 92(3):460–501.

Tesfatsion, L. (2014a). ACE Research Area: Agent-Based Macroeconomics.

Tesfatsion, L. (2014b). ACE Research Area: Learning and the Embodied Mind.

Tesfatsion, L. (2014c). Agent-Based Computational Economics: Homepage.

Tesfatsion, L. and Judd, K. L., editors (2006). *Handbook of Computational Economics: Volume 2, Agent-Based Computational Economics*. Handbooks in Economics Series, North-Holland, Elsevier.

Thaler, R. (1994). *Quasi rational economics*. Russell Sage Foundation Publications.

Tovar, C. E. (2009). DSGE models and central banks. *The Open-Access, Open Assessment E-Journal*, 3:1–31.

Trautmann, S., Vieider, F., and Wakker, P. (2009). Preference reversals for ambiguity aversion. *Tilburg Institute for Behavioral Economics Research, Tilburg University*.

Trichet, J.-C. (2010). Reflections on the nature of monetary policy non-standard measures and finance theory. *available online: http://www.ecb.int/press/key/date/2010/html/sp101118.en.html*.

Watkins, C. J. C. H. (1989). Learning from delayed rewards. *Cambridge University, Cambridge, England, Doctoral thesis*.

Woodford, M. (2003). *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press.